Observation and Modeling of Extrasolar Planets

Thesis

Observation and Modeling of Extrasolar Planets

Andrew James Carter
Department of Physics and Astronomy
The Open University

A thesis submitted for the degree of
Doctor of Philosophy
31st May 2011

Date of Submission: 31 May 2011
Date of Award: 17 May 2012

Acknowledgements

I would like to acknowledge the following people, without whom this document would either not exist or contain only the phrase "All work and no play makes Jack a dull boy." over and over in a variety of fonts and styles:

Firstly, I owe a great debt of thanks to my supervisor, Carole Haswell, for her patience, support and advice over the last four years. Having witnessed many of my fellow students struggle to subsist on stringently rationed interaction with their supervisors, I feel immensely grateful to have never lacked guidance and assistance when it was needed.

Furthermore, I would like to thank my second supervisor, Stephen Lewis; head of the Physics and Astronomy department, Nick Braithwaite; and Reader in Cosmology, Steven Sergeant for further mentorship and advice that has helped me greatly in both my research and my continuing pursuit of sanity.

A particular debt is owed to Calum MacCormick, Jonti Horner and Robin Barnard whose advice and experience was of great use to me at

the lowest ebb of my studies. Without their support and friendship, I can state with certainty that I would not have completed this thesis or the research herein.

I would like to thank my parents, Vincent and Eileen Carter, whose support has been essential over the last four years. I am exceedingly grateful and apologise that so much of said support had to be financial in nature. I would also like to thank my sister, Elizabeth, whose interest in my studies, while surprising, has always been welcome. Indeed, I would like to acknowledge all the members of my family across the globe - grandparents, aunts, uncles and cousins (and, lately, cousins-once-removeds) - all of whom I have no-doubt bored with the details of my work at some point over the last few years. Thank you, all.

I would like to acknowledge my lovely wife. But I don't have one and so can't.

To my officemates; Aga Stypczinsca, Bartoc Barc, Ros Hopwood and Yvonne Sutton, I bequeath the sum of– no, wait, wrong document... Thank you all for the company, support, assistance, patience, entertainment, pens and, most of all, a friendly ear whenever I needed to rant about something. In all seriousness, you were fantastic people with which to share four years.

I would also like to acknowledge Andrew Morris and Jonny Mason,

my housemates for a considerable portion of my studies. I owe both of you for your support, friendship and occassional maid-service. There's only one other person on the planet who I'd count among the two of you as the best friends and housemates I could have asked for - but he chose to stay in Wales, so it's his loss. I'd also like to thank Becky King and Zoe Wathen for keeping the two of them in line and quite often being defacto housemates yourselves. I'd wish the lot of you all the best for the future, but I know none of you are going to need it. Now would one set of you please hurry up and get married? It's getting ridiculous.

Further housemate-thanks goes to Ziad El Otell and, briefly, to Andrew Needham - though both of you proved to be extremely bad influences upon productivity.

Much of the programming and statitics in this thesis would be in a different state entirely were it not for the assistance of Will Clarkson whose mentorship, particularly in the development of the light-curve modeling software, was invaluable. On a related note, I would also like to acknowledge Keith Horne and David Fanning whose written works proved invaluable in the fields of statistics and programming respectively.

I would also like to thank all members of the WASP consortium, without whom this PhD would not even have begun. And, of course,

another debt of thanks is owed to all at the Space Telescope Science Institude at Johns Hopkins university for the data on which this thesis is based - we know it wasn't your fault NICMOS broke during a workshop about NICMOS (and, even if it had been, the irony alone was worth it).

Furthermore, I would like to acknowledge Ron Gilliland, who laid the groundwork for the particular strategy of observation employed herein; and both Frederick Pont and Mark Swain whose data sets form the core of two entire chapters of this thesis.

In particular, I would like to acknowledge Geoff Bradshaw for his relentless assault against the various demons plaguing the computer systems employed in the processing of the data presented herein. Whatever little remains of my sanity, I have no doubt that there would be even less were it not for Geoff's efforts. On a related note, I would also like to thank Olly Butters and Rob Farmer for timely computer-related advice at various moments of computer-related insanity.

I would not like to acknowledge Jesus. I just don't think he pulled his weight on this one.

Another particular "thank you" is directed in the vague direction of Ben Rozitis and Paul Wilkinson, whose cooperation in a mutual endeavour has been a great source of motivation in recent months. Ben, congratulations on finishing first. Paul... Sorry, but someone

had to lose. (And a further apology to Jo Iacovides - but there's no use complaining; you knew what you were agreeing to.)

There are many friends whose advice and encouragement have been invaluable over the course of my research. While there are far too many to mention, I would like to make particular note of Tom Williams who, for unfathomable reasons, was the first person to request a copy of this document upon its completion. A further acknlowledgement goes to Tom's lovely wife, Alice, and son, Ioan, who will find themselves responsible for Tom's wellbeing once the contents of said document has robbed him of his higher brain functions, as it has mine.

Another special thanks goes to Becky and Marcus Fitzsimons; a couple whom I was honoured to serve as best-man during my studies. I owe both of you, and indeed all of my friends back home in Yorkshire, a huge apology for finding so little time to visit you all over the last few years. It is my hope that, upon reading this document, you might finally understand why that happened.

One special acknowledgement goes to Jonty Marshall, Andy Mason and Vicky Bending for managing the supply of tea, coffee, biscuits and doughnuts essential to the construction of this document.

I would also like to acknowledge the contributions of Richard Busuttil, who supplied the Roche lobe plots found in Chapters 7 and 8; of Anthony Davenport, Jon Dawson, Maria Duffy, Sonia Newman

and Danny Barthaud, whose contributions are entirely fictional; and of Liam Steele, whose contribution is purely hypothetical and likely incorrect.

Also worthy of note are (almost in order of clicking randomly in the list) Rebecca Wolsey, Neil Murray, Naomi Murdoch, Katarina Milijovic, Ben Dryer, Stacey Birket, Sunitha Pangala, Paul Davey, Marc Scully, James 'Hero' Tutt, Kris Froud, Euan Monaghan, Zdenek Masin, Ania Kowalczyk, Chris Cottis, Kathryn McDermott, Claire Parfitt, Paul Steele, Jemma Davidson, Dan Fawcett, Liz Uttley, Lana George, Richard 'Danger' Harriss, Chris Mansell, Lindsey Shaw Greening, Adam Rae, Samantha Harrison, Stephan and James Holmes (no relation), Annika Simpson, Min Tran, Kuba Bochinski, Rebecca Wilson, Neil Parley, Brian Pluss, Phil Davis, Elena Nickson, Pete Landsberg, Feargus Abernethy, Bethany Alden, Sophie Allan, Jonathan Keelan, Tim Tomkinson and the ever-cheerful Andreas Morlok - all of whom contributed significantly to most of the delays to this document.

I should also thank every postgraduate student, post-doc, academic, miscelaneous Open University staff member and random stranger I have encountered during my studies. I would attempt to name all of you, but that would innevitably end with me missing someone out - and that is a risk I'm not willing to take. Therefore, I include the following get-out clause:

A very special Thank You to [name redacted to avoid causing embarassment, you know who you are] - just for being you.

Finally, of my friends, I would like to acknowledge the writers I have come to know over the last two years - particularly Hina Malik, Alli Parker, Lucy Hay, Andrea Mann, Frank Brinkley, Mina Zaher and Anton Russel. The reminder of the future that awaits me once my studies are complete has been all the more inspiring thanks to you.

I would also like to thank *you*, dear reader, for inexplicably taking the time to read this thesis - or perhaps even just to leaf through these acknowledgements. Maybe you're a PhD student yourself - probably one I will never meet. Maybe you're looking for words of wisdom in these pages to help you with your own studies - in which case: Get your own wisdom, you lazy sod. This is mine - you can't have it.

Finally, I would like to sarcastically thank whichever God thought it would be funny to have me bludgeoned unconscious by a cricket ball in my second year. You're hilarious.

Abstract

The field of exoplanet research has currently yielded the discovery of 552 planets. This figure includes 132 transiting planets which can be studied in greater detail and have formed the cornerstone of research to characterise the exoplanet population. In particular, such studies seek to analyse the planetary atmospheres, but research has thus far yielded more questions than answers.

Exoplanetary atmospheric studies have typically focussed on one planet apiece - complicating any comparative analysis as every result employs different methods and instruments. For a comprehensive, comparative study, a robust and reliable means of reducing and analysing such observations is required, along with a body of data from a single instrument.

One such instrument is the Hubble Space Telescope (HST) whose NICMOS (Near Infrared Camera and Multi-Object Spectrometer) instrument has observed the transits of nine extrasolar planets across multiple wavelengths in the near-infrared. A robust pipeline has been developed to reduce all such observations using the same techniques.

This pipeline reduces grism images of an exoplanet host star across a transit event. These exposures are checked for bad pixels, flat fielded and background-subtracted before robust extraction of a transit light curve. This light curve is then detrended to remove systematic noise by application of a new technique devloped in this study. Following detrending, the light curve is modelled using a bespoke MCMC (Markov-Chain Monte-Carlo) algorithm to determine the planetary parameters. A continuum of wavelength-dependent transit light curves is also extracted, detrended and modelled to determine the variation in transit depth with wavelength; and thereby infer the transmission spectrum of the planet's atmosphere.

The finished pipeline has been applied to three sets of HST NICMOS observations covering the transits of WASP-2b, HD189733b and GJ436b. For each data set, a new set of planetary parameters has been derived and for WASP-2b and HD189733b an atmospheric transmission spectrum extracted. Both spectra show signs of atmospheric haze and molecular absorption, but also evidence of residual systematic noise, complicating analysis.

Contents

Chapter 1

Introduction

1.1 Chapter Outline

The following chapter comprises a summary of the field of exoplanetary research with respect to the research conducted herein. An overview of the origins of extrasolar planetary science is followed by a review of the main detection techniques - in particular the Radial Velocity and Transit techniques, but also covering Transit Timing, Microlensing, Direct Detection and Pulsar Timing - detailing their methodology, associated science (e.g. Limb Darkening), advantages and drawbacks. The chapter concludes with details on the currently unanswered questions in the field; the density problem, atmospheric analysis and planet formation.

1.2 A History of Extrasolar Planets

Since the realisation that each star in the sky was a sun, much like our own, mankind has theorised on the nature of the planets which might be found there. However, we would have to wait until 1992 for the first evidence that such planets existed (Wolszczan and Frail, 1992, see Section 1.3.5). After that, the first

detection of a planet around a solar-type star came in 1995 (Mayor and Queloz, 1995) and spawned a whole new field of research into the nature of the extrasolar planets.

The total number of extrasolar planets (also called 'exoplanets') now exceeds 500 (Schneider, 2010) and is growing at such a rate as to ensure any more exact figure would become out of date in short order. The majority of these planets are 'hot Jupiters'; planets of similar size and mass to Jupiter, but which orbit very close to their parent stars and on very short timescales. It is not presumed that the majority of planets are Hot Jupiters (Mordasini et al., 2009); their current prevalence is most likely a detection bias resulting from current limits on the precision of modern instrumentation (see Section 1.3).

However, while these hot Jupiters do not tell the full story of the population of planets, they offer valuable insight into theories of planet formation (Alibert et al., 2010; Maciejewski et al., 2010a, see Section 1.4.3) and, more recently, research has begun shedding light onto the atmospheric composition of these planets (see Section 1.4.2).

As technology improves and new exoplanet searches begin, the range and population of known exoplanets continues to expand. At time of writing, the largest known exoplanet, TrES-4b (Mandushev et al., 2007), stands at $1.81 R_J$ (Jupiter Radii, $1 R_J \simeq 11.21 R_\oplus \simeq 7.15 \times 10^{11} m$) with a mass of $0.877 M_J$ (Jupiter Masses, $1 M_J \simeq 317.8 M_\oplus \simeq 1.90 \times 10^{27} kg$), compared to the most massive planet, CoRoT-3b (Deleuil et al., 2008), at $21.66 M_J$ with a radius of $1.01 R_J$ (see Section 1.3.2.5).

Comparatively, the smallest known planet (around a solar-type star), Kepler-9d (Holman et al., 2010), has a radius of just $0.147R_J$ (or $1.64R_\oplus$) and a mass currently estimated at $0.022M_J$ (or $6.99M_\oplus$). GJ581e (Mayor et al., 2009), with a mass of just $0.006M_J$ ($1.94M_\oplus$) may be smaller still, but its radius is currently unknown (see Section 1.3.1).

The small semi-major axes of hot Jupiters' orbits result in very short orbital periods with the shortest, WASP-19b (Hebb et al., 2010), orbiting its parent star in 0.789 days (less than 19 hours) at a distance of just 0.0164AU (Astronomical Units, $1AU = 1.496\times10^{11}m$). However, as the field of known exoplanets increases, the number of more distant bodies is growing. HR8799b (Marois et al., 2008) is currently the furthest known planet from its star for which both the period and mass are known; with an orbital separation of $68AU$ and a period of 1.7×10^5 days ($\sim$ 466 years).

Exoplanet research is beginning to reveal that extrasolar systems have a great deal of variety. At time of writing there are 52 confirmed systems with multiple planets[1] (Schneider, 2010) ranging from systems with several Hot Jupiter type planets, such as 55 Cancri (Fischer et al., 2008), to those not too dissimilar to our own Solar System, like HD40307 (Mayor et al., 2008). The larger the sample of known exoplanets, the better our understanding of the population as a whole becomes.

[1]Not counting the hundreds of likely candidates discovered by Kepler (Borucki et al., 2011).

1.3 Detection Techniques

Various methods have emerged as effective means of detecting extrasolar planets. As technology improves and the field of research widens, those methods become more refined and less affected by various detection biases while, simultaneously, new methods arise to discover new populations of planets and improve our understanding as a whole.

However, at time of writing, only a handful of such techniques have proven themselves indispensible in the field of exoplanet detection - most specifically the Radial Velocity and Transit techniques (see Sections 1.3.1 and 1.3.2 respectively) which, between them, account for over 92% of the known exoplanet population (Schneider, 2010).

1.3.1 Radial Velocity

The radial velocity technique was the first method successful at detecting planets around solar-type stars (Mayor and Queloz, 1995) and has since detected over 450 exoplanets in almost 400 systems (Schneider, 2010). It is, by any definition, the most successful exoplanet detection technique yet implemented, but not necessarily the most useful, as will be discussed below.

The technique works by measuring the radial velocity of the parent star in any exoplanetary system by examination of its spectral lines (Sasselov, 2008). In any planetary system, the star and planets orbit a common centre of mass which (excluding multiple star systems) is typically located within the star itself. The result is that the star itself moves on a tightly confined orbital path - effectively

appearing to 'wobble'. The magnitude and period of this orbit is directly related to the mass and period of the planets in the system and so, by measuring the stellar wobble, it is possible to infer an upper limit on the mass and orbits of the system's planets.

The wobble itself is measured from the movement of various absorption lines in the star's spectrum. As the star moves about its orbit, some component of its motion will be in parallel to our line of sight - the portion of the movement directly towards or away from us. This motion causes a Doppler shift in the frequency of the observed light; blueshifting the spectral lines as the star moves towards us and redshifting them as it moves away.

The magnitude of this shift is dependent on both the mass of the planet (relative to the star) and the semi-major axis of the planet's orbit. For example, gravitational pull of the Earth moves our own sun at a velocity[1] less than $0.1ms^{-1}$. By comparison, if the Earth had a mass comparable with Jupiter, that velocity shift would increase to almost $30ms^{-1}$.

For an exoplanet, the size of the Doppler shift is measured in comparison to a static source at the point of observation. By passing the detected light through a known medium (typically iodine, Konacki et al. (2005)) the spectrum of that medium is imprinted on top of the stellar spectrum. Then, as the star's spectral lines vary in wavelength due to the Doppler shift, the imprinted lines remain stationary; providing a baseline against which the magnitude and period of the

[1]Assuming an observer on the ecliptic plane whose position is static - i.e. independent of both star and planet.

Doppler shift can be measured. From this measurement, the radial velocity of the star can be infered from:

$$\frac{\lambda_0}{\lambda} = \left(1 - \frac{v_S}{c}\right) \tag{1.1}$$

where λ_0 is the rest wavelength of any given line in the stellar emission spectra; λ is the measured wavelength of that line; v_S is the relative velocity of the target star in relation to the Earth; and c is the speed of light.

1.3.1.1 Calculating Planetary Properties

Measuring the relative velocity, v_S, of the star over a complete orbit of any exoplanet[1] produces a sine curve like that seen in Figure 1.1. The period of the sine curve is equal to the period of the planet's orbit and, through Kepler's third law, allows us to calculate the separation between the planet and the star as:

$$a^3 = \frac{GM_S}{4\pi^2}P^2 \tag{1.2}$$

where a is the semi-major axis of the planet's orbit; G is the gravitational constant; M_S is the mass of the star; and P is the period of the planet's orbit.

The mass of the parent star, M_S, can be estimated from its luminosity and spectral type, assuming the star is on the main sequence. Exoplanets have been

[1] Assuming said planet is the only other massive body in the system and follows an ostensibly circular orbit.

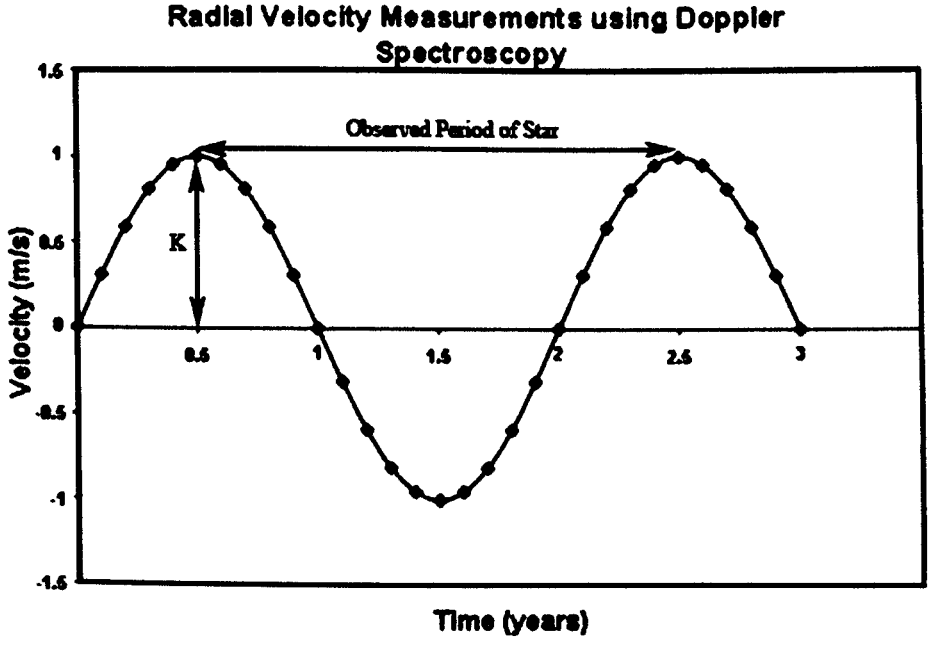


Figure 1.1: *The change in measured relative radial velocity, v, of a star resulting from the orbit of its planet. The overplotted sine curve demonstrates the best-fit motion of the curve resulting from the star orbitting the common center of mass with.*

found orbiting non-main sequence stars (for example: PSR1257 + 12 (Wolszczan and Frail, 1992) or HD13189 (Hatzes et al., 2005)) and in such cases alternate models and methods exist for estimating the stellar mass. However, for the purposes of this research, we will consider only planets orbiting main sequence stars.

Once the orbital separation, a, has been calculated, it is possible to derive a lower-limit on the mass of the planet itself, M_P. This is calculated from the amplitude, K, of the sine curve in Figure 1.1. By considering the relative orbital velocities of the star and planet and assuming the angular momentum of the system is conserved (i.e. that there are no other bodies present beyond the star

and planet), we can state that:

$$M_S v_S = M_P v_P \tag{1.3}$$

where M_S is the mass of the star; v_S is the orbital radial velocity of the star; M_P is the mass of the planet; and v_P is the orbital radial velocity of the planet.

The orbital velocity of the star, v_S can be calculated from the amplitude of the sine curve, K. Since K is the relative velocity of the star to the barycentre of the extrasolar system parallel to our line of sight, the actual velocity of the star can be stated as:

$$v_S = \frac{K}{\sin i} \tag{1.4}$$

where v_S is the orbital radial velocity of the star; K is the velocity of the star parallel to our line of sight (also the amplitude of the sine curve in Figure 1.1); and i is the inclination of the orbital plane of the extrasolar system perpendicular to our line of sight, as demonstrated in Figure 1.2.

The orbital velocity of the planet, v_P, can be calculated from the known orbital separation of the star and planet (see Equation 1.3.1.1) as:

$$v_P(t) = \sqrt{\frac{GM_S}{r(t)}} \tag{1.5}$$

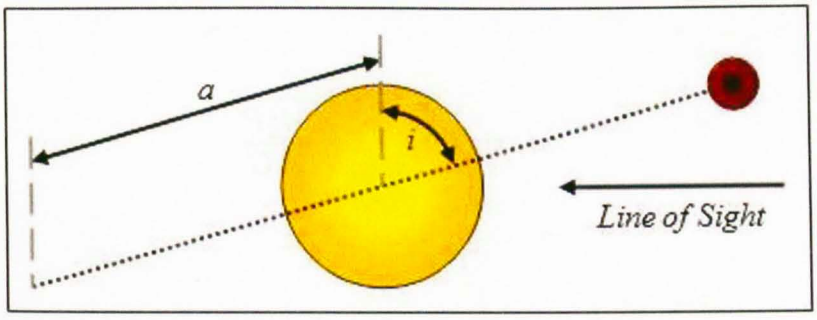


Figure 1.2: *The inclination of an exoplanet's orbit (with semi-major axis, a) relative to our line of sight. The angle of inclination, i, is measured perpendicular to our line of sight so that a planet which transits (see Section 1.3.2) directly across the centre of its parent star would have an inclination angle of 90°. From transit observations, it is not possible to discern if the system's inclination is above or below our line of sight; therefore, the inclination angle is always stated as a figure $\leq 90°$.*

where $v_P(t)$ is the orbital velocity of the planet at a time, t; G is the gravitational constant; M_S is the mass of the star; and $r(t)$ is the orbital separation between the star and planet at a time, t. In a system with a perfectly circular orbit, the orbital separation (r) is constant and so can be equated directly to the semi-major axis of the system (a), as calculated in Equation 1.3.1.1.

Therefore, we can calculate the mass of the planet, M_P as:

$$M_P = \frac{K}{\sin i}\sqrt{\frac{aM_S}{G}} \tag{1.6}$$

Or, alternatively, as:

$$M_P \sin i = K \sqrt[3]{\frac{M_S^2 P}{2\pi G}} \tag{1.7}$$

where M_P is the mass of the planet; K is the radial velocity of the star parallel with our line of sight; i is the inclination angle of the system; a is the semi-major axis of the planet's orbit; P is the period of that orbit; M_S is the mass of the star; and G is the gravitational constant.

In situations where the planetary orbit is non-circular, these calculations become more complex - factoring in the eccentricity of the planet's orbit which can also be measured from the radial velocity curve. The impact of eccentricity on the observed Doppler shift is dependent both on the scale of eccentricity and the orientation of the orbit's semi-major axis in relation to our line of sight.

For example, as figure 1.3 demonstrates, the curve produced by a planet whose semi-major axis is parallel to our line of sight becomes skewed - with either the rise or fall of the curve contracted at perihelion (at the planet's maximum orbital velocity) and, conversely, the fall or rise of the curve is extended at aphelion. For a system whose semi-major axis is perpendicular to our line-of sight, the peri- and aphelion correspond with the peaks and troughs of the radial velocity curve (or vice-versa) and result in a sharpening (at perihelion) and flattening (at aphelion) of these features.

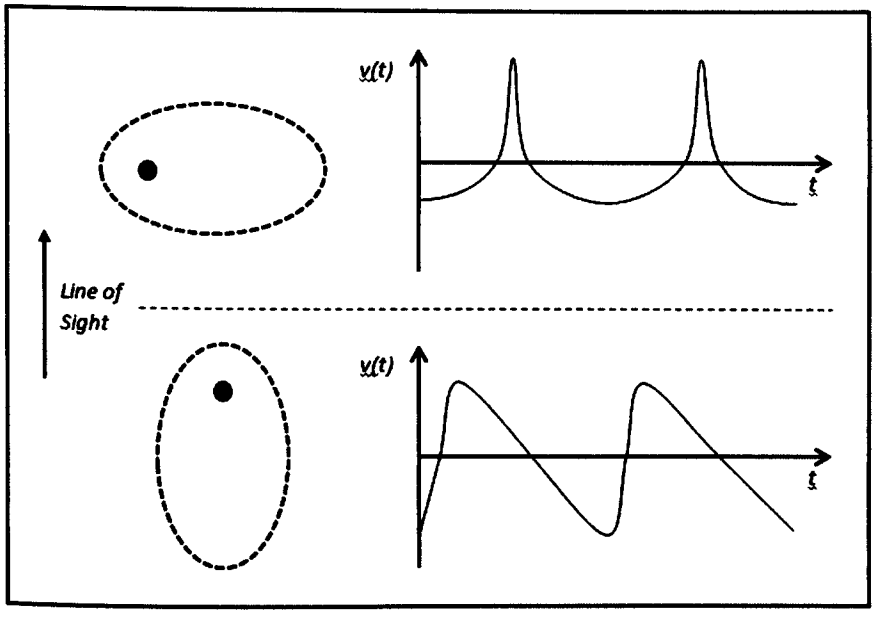


Figure 1.3: *The radial velocity curves (right) produced by a planet with an eccentric orbit whose semi-major axis is perpendicular (top) and parallel (bottom) to our line of sight.*

1.3.1.2 Drawbacks and Detection Bias

Note that the inclination angle of the system, i, is not measured by the radial velocity detection technique. As a result, the only mass which can be put to the planet is a lower limit - based on the assumption that the orbital plane of the system is parallel to our line of sight; thereby meaning we see the full motion of the star ($i = 90°$, $K = v_S$) and not just some component of it. This lower limit is typically stated as $M_P \sin i$.

The radial velocity curve (Figure 1.1) makes this detection technique especially sensitive to massive planets with small orbital radii and periods (i.e. hot Jupiters). Such planets exert a greater gravitational pull on their parent star and so result in a sine curve with a greater amplitude - making it much easier to detect. Furthermore, the short orbital periods of such planets mean that a detection can be verified relatively quickly, compared to planets whose orbits last months or years.

Indeed, at the birth of the field of exoplanet research, telescope technology was insufficiently advanced to even hope to detect planets with masses comparable to the Earth. While some instruments had the capability, the chance of detecting a planet around any given star was (and still is) largely unknown - thereby making anything short of a wide-area survey a grossly inefficient and optimistic use of telescope time.

However, as technology has improved and a wealth of knowledge and experience gathered, the limits of detection have been pushed back, bringing us ever closer to discovering the first true Earth-analogues orbiting other stars. As it

stands, that detection is likely to happen very soon and, at present, the radial velocity technique remains the benchmark for verifying all exoplanet discoveries - though, as other techniques exceed the limits of radial velocity, alternative confirmation methods (such as transit timing Lissauer et al. (2011)) are being sought. However, determining more detail about the planet's parameters and properties requires other methods to be implemented.

1.3.2 Transits

An extrasolar planetary transit occurs when an exoplanet passes directly between our line of sight and its parent star. During this event, the planet obscurs a small fraction of light from its star, creating a characteristic dip in the recorded light - the transit light curve (see Figure 1.4).

The first exoplanetary transit was detected in 1999 by Henry et al. (2000). The existence of HD 209458 b was already known, having been detected by radial velocity measurements (see Section 1.3.1) conducted by the Keck telescope in Hawaii. However, this first detection of an exoplanetary transit allowed for a far greater wealth of information to be gathered on the planet.

The first planet detected by its transit alone was OGLE-TR-56b in 2002 (Udalski et al., 2002) and, since then, over 100 planets have been discovered by various transit detection surveys and confirmed by follow-up radial velocity observations. Such follow-up observations are required because of various other astro-phenomena capable of mimicking a transit light curve.

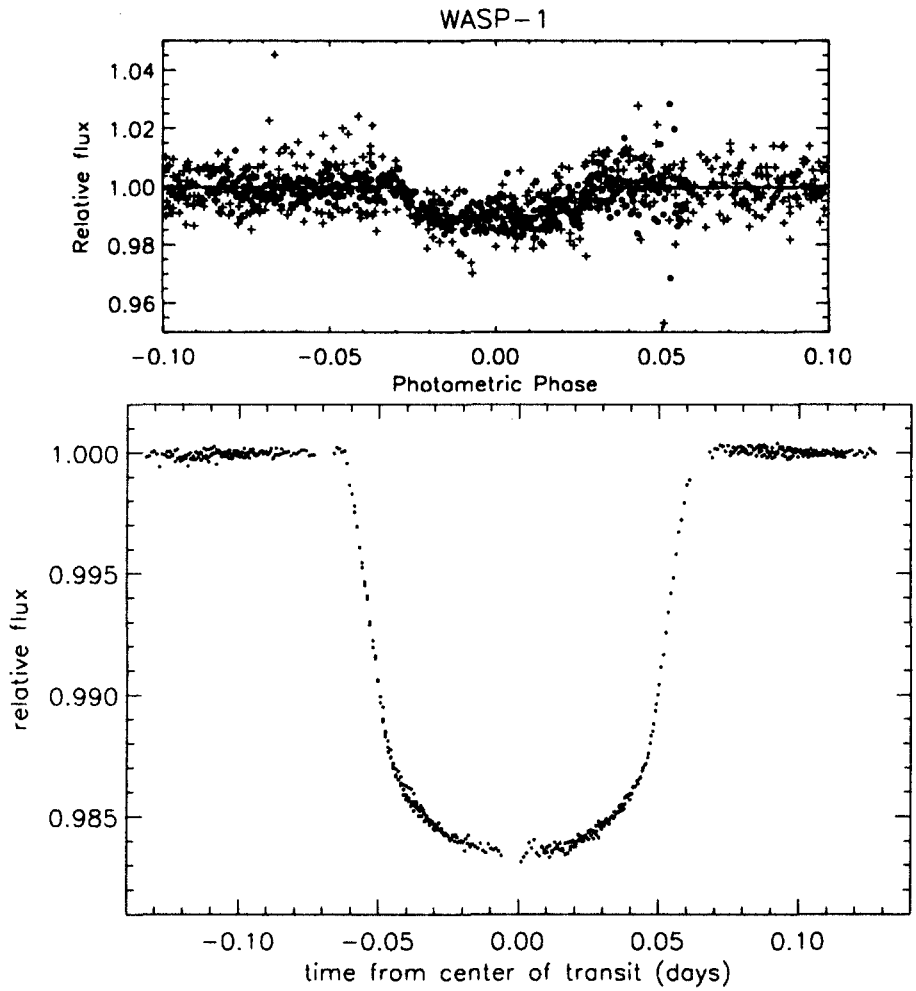


Figure 1.4: *The low-S/N transit light curve of WASP-1b (top), formed as the planet passes across the disk of its parent star, eclipsing approximately 1% of the star's light (Cameron et al., 2007). In contrast, the high-S/N light curve (bottom) shows the clear U-shaped transit of HD 209458 b.*

1.3.2.1 Calculating Planetary Properties

Unlike those planets whose presence is known through the radial velocity alone, transiting planets reveal much more information than can be derived from other means. The transit light curve itself has a distinct shape (see Figure 1.4) which can be divided into sections according to the aspects of the transit itself. The divisions between these sections are known as the 'contact points' of the transit, also termed C_1, C_2, C_3 and C_4.

The first such contact point (C_1) occurs when the limb of the planet meets the limb of the star at the start of the transit and the beginning of the section of the light curve known as the 'ingress'. The second contact point (C_2) occurs when the whole of the disk of the planet has moved over the star (i.e. having covered an orbital distance of one planetary diameter, or $2R_P$, since C_1). Between C_2 and C_3 is the 'transit floor' where the disk of the planet remains wholly over the disk of the star. Once the limb of the planet passes the limb of the star (C_3), the planet enters the 'egress' phase of the transit which continues until the planet has moved wholly from the disk of the star (C_4) at the end of the transit event.

The planet is considered 'in-transit' at any time between C_1 and C_4. At all other times, the system is considered out-of-transit (sometimes abbreviated to OOT). The time at which each transit can be expected to occur can be calculated from the transit's ephemeris and the period of its orbit (P). The various sections of the transit light curve are also demonstrated in Figure 1.5.

The precise shape of the transit light curve is dependent on various factors; most significantly the size of the planet itself. The depth of transit allows the

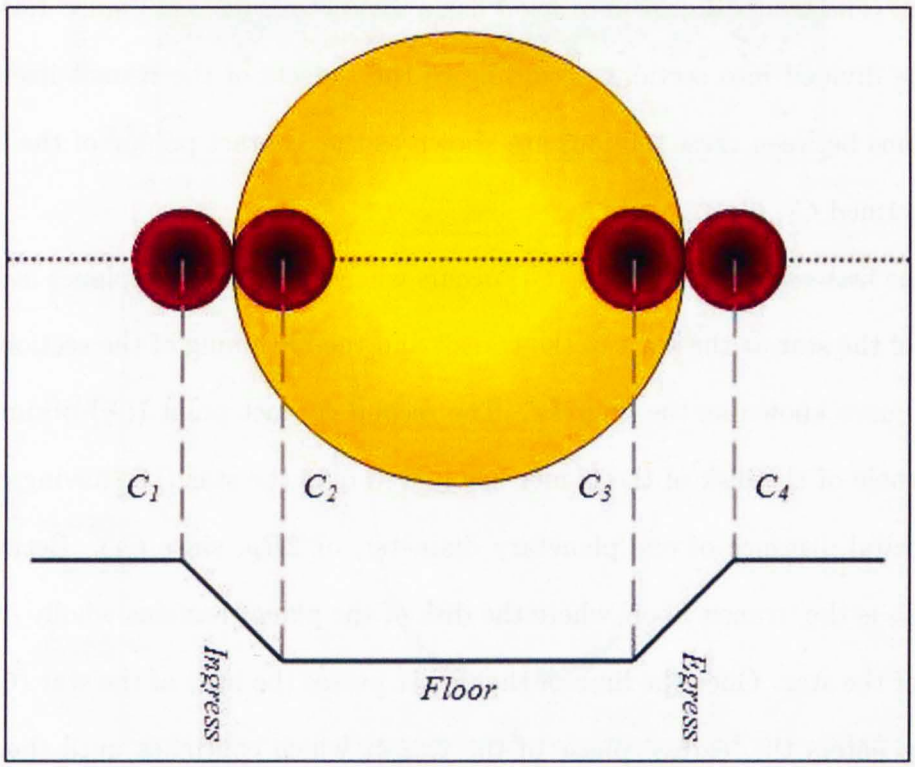


Figure 1.5: *The contact points of an exoplanetary transit. As the planet transit its star from left to right, the transit ingress begins at C_1 when the planet begins to move over the disk of the star and ends at C_2 once the whole of the planet has moved over the stellar limb. The transit floor then continues until C_3 when the planet begins to move off the disk of the star. This begins the transit ingress which ends at C_4 with the planet moving wholly off the stellar disk at the end of the transit event.*

planet's radius (R_P) to be measured, something not currently possible with other techniques. If we consider, for the sake of simplicity, that the disk of the star can be approximated to a circle of uniform intensity and the disc of the planet to a smaller, completely opaque disc, then the recorded flux from the star can be approximated as:

$$F = IA \tag{1.8}$$

$$F_{OOT} = I\pi R_S^2 \tag{1.9}$$

$$F_{Floor} = I(\pi R_S^2 - \pi R_P^2) \tag{1.10}$$

where F is the observed flux from the star; I is the intensity of light from the star per unit area; A is the observable area of the disk of the star; F_{OOT} is the recorded flux while the planet is out of transit; F_{Floor} is the recorded flux while the planet is wholly in-transit (i.e. at the transit floor); R_S is the radius of the star; and R_P is the radius of the planet.

However, it is easier to measure the flux from the parent star as a percentage change over time, rather than a precise value. Therefore, we can express the above relationships as:

$$\left(\frac{\Delta F}{F_{OOT}}\right) = \frac{F_{OOT} - F_{Floor}}{F_{OOT}} = \frac{L\pi R_S^2 - L(\pi R_S^2 - \pi R_P^2)}{L\pi R_S^2} = \left(\frac{R_S}{R_P}\right)^2 \tag{1.11}$$

And so we can determine the radius of the planet to be:

$$R_P = R_S\sqrt{\frac{\Delta F}{F_{OOT}}} \tag{1.12}$$

where R_P is the radius of the transiting planet; R_S is the radius of the star; and $\Delta F/F_{OOT}$ is the depth of transit.

Furthermore, analysis of either the ingress and egress of the transit or the transit duration allows for the calculation of the inclination angle of the system (i). This in turn allows the actual mass of the planet (M_P) to be calculated from the known lower limit ($M_P \sin i$) imposed by radial velocity observations (see Section 1.3.1).

As demonstrated in Figure 1.6, the duration of the transit is dependent on the distance the planet must travel across the disk of the star - a figure which diminishes along with the inclination angle of the system (an inclination angle of 90° would result in a transit of maximum possible duration). As the inclination angle decreases, the path of the planet across the star moves further towards the star's limb. This difference is characterised by the distance between the centres of the star and planet at closest-approach; known as the 'Impact Parameter':

$$b = a\cos i \tag{1.13}$$

where b is the impact parameter of the transit; a is the semi-major axis of the planet's orbit; and i is the inclination angle of the system.

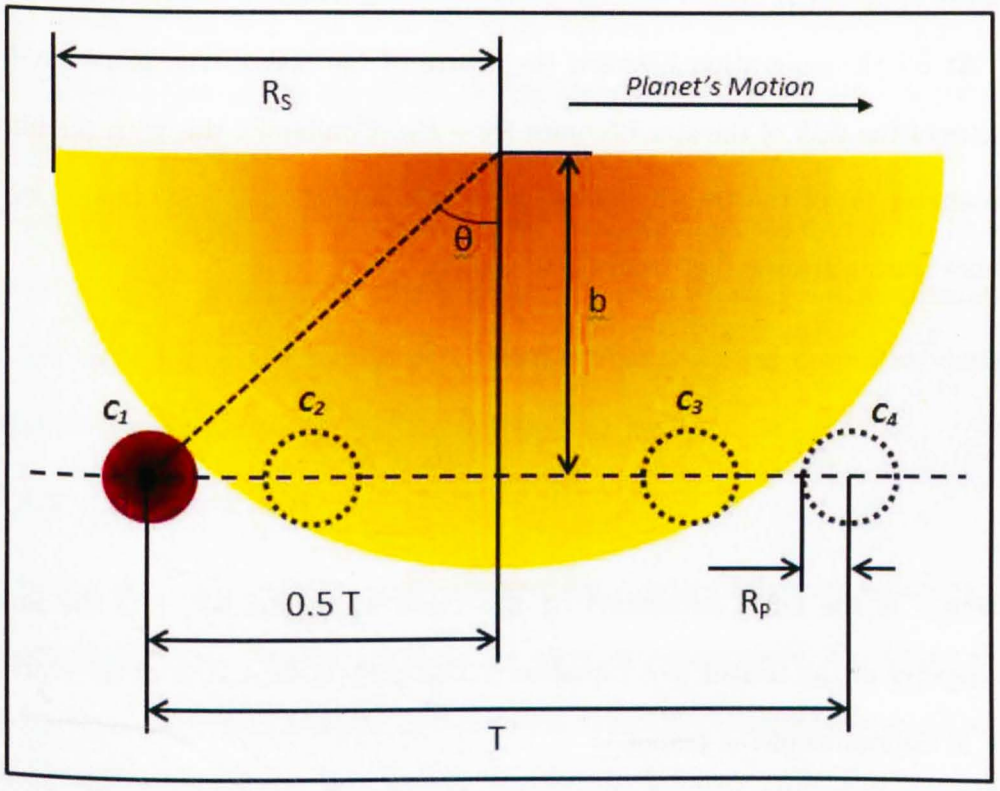


Figure 1.6: *The geometry of a transit event at C_1 describing the duration of the transit over the distance T through the angle 2θ with the centre of the disk of the star.*

The path of the planet during transit is a chord across the disk of the star (extended from the limb of the star by one planetary radius) and, as such, can be characterised by the angle it subtends with the centre of that disk. For simplicity's sake, we can consider this in terms of the angle subtended between C_1 and the transit midpoint provided we double the resulting distance calculations.

At C_1 the separation between the centre of the disk of the planet and the centre of the disk of the star becomes $R_S + R_P$. Combining this with the impact parameter (b) of the transit, allows us to calculate the angle subtended by the planet during transit (i.e. from C_1 to C_4) as:

$$2\theta = \cos^{-1}\left(\frac{b}{R_S + R_P}\right) \tag{1.14}$$

where θ is the angle subtended by the chord of the transit; b is the impact parameter of the transit (see Equation 1.3.2.1); R_S is the radius of the star; and R_P is the radius of the planet.

From this we can calculate the length of the transit chord as:

$$D_T = 2(R_S + R_P)\sin\theta \tag{1.15}$$

Combining this with the orbital period allows us to define the transit duration as:

$$t_T = P\frac{2R_S \sin\theta}{\pi a^2} = P\frac{D_T}{\pi a^2} \tag{1.16}$$

where t_T is the duration of the transit; P is the period of the planet's orbit; R_S is the radius of the star; and 2θ is the angle subtended by the transit chord; D_T is the distance traveled by the planet during transit; and a is the semi-major axis of that planet's orbit.

As such, Equation 1.3.2.1 allows for the inclination of the system to be calculated from the period of the planet's orbit (P, from which a is also calculated) and the duration of the transit (t_T). Further discussion on properties derived from transit light curves can be found in Section 1.4.2.

1.3.2.2 Temperature

An important parameter in exoplanet research, particularly in relation to the planet's density (see Section 1.4.1), is the planet's temperature, T_P. Given their proximity to their parent stars, hot Jupiters typically maintain an equilibrium temperature upwards of 1000 Kelvin due to the incident radiation from their parent stars. If we measure the luminosity of the star, we can determine the flux of radiation emitted as:

$$F_S = \frac{L_S}{4\pi r^2} \tag{1.17}$$

where F_S is the radiated flux from the star (energy per unit area per unit time); L_S is the luminosity of the star; and r is the radial distance from the star at

which the flux is measured. In the case of a planet, r would be the semi-major axis, a.

Applying this to the disk of the planet, we can determine the energy of the incident radiation as:

$$\dot{E}_I = \pi R_P^2 F_S = \frac{L_S R_P^2}{4a^2} \tag{1.18}$$

where $\dot{E}_I$ is the incident energy radiated onto the planet per unit time; R_P is the radius of the planet; F_S is the flux from the star; L_S is the luminosity of the star; and a is the planet's orbital semi-major axis.

If we assume the planet to be in thermal equilibrium[1] then the incident energy must be equalled by the energy radiated by the planet.

Some portion of this energy would be the reflected starlight - some portion of $\dot{E}$ not absorbed by the atmosphere. The exact proportion of reflected light is dependent on the planet's albedo, A, which is difficult to measure precisely. However, at such high temperatures, it is likely that the albedo of hot Jupiters is very low - this allows us to still place an upper estimate on the tempterature of the planet, T_{Max} which is unlikely to be much higher than the planet's actual temperature and is still applicable when comparing an exoplanet to the population as a whole.

[1]The rotation of any exoplanet should be sufficient to ensure its surface reaches thermal equilibrium as each part of its atmosphere spends the same time in the planet's shadow and the same time in the hemisphere irradiated by the star - allowing the temperature to be evenly distributed across both hemispheres. However, planets with a very small semi-major axis (a) may be tidally locked to the star. In this instance, the same hemisphere would face the star at all times - likely resulting in a large temperature gradient between the dayside and nightside of the planet.

If we approximate the planet's thermal radiation to a black body, we can state that:

$$\dot{E}_R = 4\pi R_P^2 \sigma_B T_{Max}^4 \tag{1.19}$$

where $\dot{E}_R$ is the energy radiated by the planet as a black body; R_P^2 is the planet's radius; σ_B is the Stefan-Boltzman constant ($5.6704 \times 10^8 Wm^{-2}K^{-4}$); and T_{Max} is the upper limit on the planet's Equilibrium temperature.

Therefore, we can estimate the planet's temperature by equating $\dot{E}_I$ to $\dot{E}_R$, such that:

$$\frac{L_S R_P^2}{4a^2} = 4\pi R_P^2 \sigma_B T_{Max}^4 \tag{1.20}$$

Or:

$$T_{Max} = \left(\frac{L_S}{16\pi\sigma_B a^2} \right)^{1/4} \tag{1.21}$$

where T_{Max} is our upper-limit on the planet's equilibrium, black-body temperature; L_S is the luminosity of the star; and a is the planet's semi-major axis.

If the albedo of the planet is known, the exact temperature can be calculated from T_{Max} by:

$$T = T_{Max}(1 - A)^{1/4} \tag{1.22}$$

where T is the planets equilibrium, black-body temperature; T_{Max} is our previously calculated upper-limit on that temperature; and A is the planet's albedo. However, with the value of A likely to be small for hot Jupiters, the factor of $(1 - A)^{1/4}$ is not likely to make a large difference to the resulting temperature.

1.3.2.3 Limb Darkening

Thus far we have assumed the disk of the star to be a circle of uniform luminosity. However, in reality, the stellar disk appears brighter at its center than it does at its limb; an effect known as 'limb darkening'.

Light emitted at the centre of the star must travel outwards through the stellar atmosphere before it escapes the star. When observing the disk of the star, we observe only those photons which escape the stellar atmosphere and travel along our line of sight. Therefore, photons emitted from the same layer of the stellar atmosphere from the limb of the disk star will travel a greater distance through the atmosphere than those emitted at the centre.

This greater path-length through the stellar atmosphere increases the probability that the photon will be scattered by the plasma of the stellar atmosphere, to a degree dependent on the optical depth. As a result, fewer photons from deeper within the atmosphere escape the star travelling along our line-of-sight as we move closer to the star's limb.

While there is no precise law describing limb darkening, there are various models which approximate its effect. Transit studies typically adopt the following quadratic law to model the impact of Limb Darkening on the light curve:

$$\frac{I(\gamma)}{I} = 1 - U_0(1 - \cos\gamma) - U_1(1 - \cos\gamma)^2 \quad (1.23)$$

where $I(\gamma)$ is the intensity of light emitted from a point on the disk of the star whose radius from the centre of the star subtends an angle, γ, from our line of sight (see Figure 1.7); I is the intensity of light emitted at $\gamma = 0$ (i.e. at the center of the disk of the star); U_0 is the linear coefficient of the limb darkening model and U_1 is the quadratic component of the model (Haswell, 2010; Mandel and Agol, 2002).

The values of U_0 and U_1 can then be used to define the intensity at any point on the disk of the star. Alternate models of limb darkening are available, including both cubic and logarithmic approximations, but such complexity would be detrimental to this research. The effect of limb darkening is expected to be small in the infra-red and a new model will need to be generated for each star. Therefore, the ideal limb darkening model in this context is one flexible enough for a transit model based upon it to produce a reliable fit to the measured light curve while also being simple enough to require minimal processing time during the simulation of such a model. As such, a quadratic approximation to limb-darkening is the best choice for the purposes of this research.

Since the optical depth of the stellar atmosphere is wavelength dependent, the limb darkening of a star will also vary with wavelength. As a result, any limb darkening coefficients ascribed to a star will be accurate only for a certain

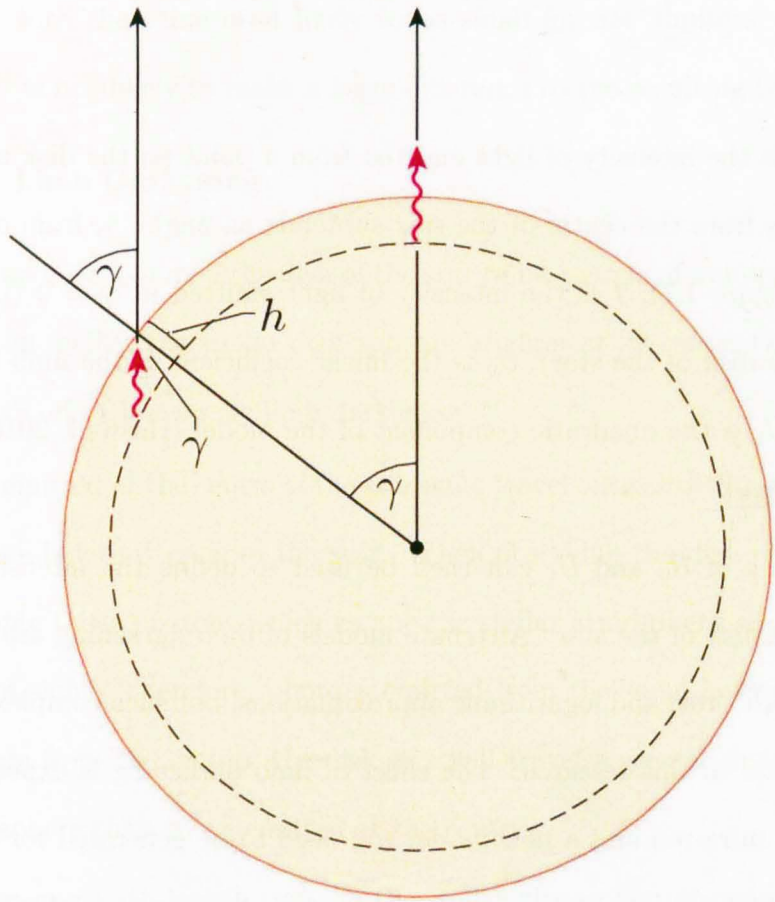


Figure 1.7: *Two photons (red arrows) emitted from the same depth into the stellar atmosphere (h), both travelling along our line of sight. The photon emitted from the centre of the stellar disk (from our viewpoint, top) travels a distance of h through the atmosphere. In contrast, the photon emitted towards the limb of the star (left) travels a greater distance of $\sim (h/\cos\gamma)$. As a result, the photon towards the edge of the disk is more likely to be scattered. Applying this to all photons emitted by the star, a smaller proportion of photons from any given depth will escape the star at the limb compared to the centre.*

wavelength range. Thankfully, this variation is small across the infra-red range covered by the data sets used in this research (see Chapter 2) and, while this does not necessarily mean its impact will be too small to produce measureable effects, it is not so great as to prevent one set of initial estimates being applied across the full wavelength range of any given data set.

Overall, the shape of the transit light curve can be said to be dependent on the ratio of radii between the planet and star (R_P/R_S), the inclination angle of the system (i), the semi-major axis of the planet's orbit (a), the period of the transit (T) and the two limb darkening coefficients for our quadratic model (U_0 and U_1). The precise impact of these parameters on the shape of the transit will be discussed further, with reference to modeling such light curves, in Chapter 5.

1.3.2.4 False Positives

As has been previously noted, follow-up observations of transiting planet candidates are essential to rule out various scenarios capable of mimicking an exoplanetary transit.

The most common of these false positives is a grazing eclipsing binary star system. In this instance, the two stars in a binary system execute a grazing eclipse twice during each complete orbit. In each grazing eclipse, one star obscurs part of the second star's limb; blocking a small portion of the total light and producing a resulting light curve similar to that of a transit.

Grazing eclipses produce a much sharper, V-shaped 'transit' than would be observed for an exoplanet and would typically produce two transit-like events

(assuming a low orbital eccentricity). In most cases, the magnitude of the stars would be disparate enough for these 'transits' to have noteably different depths, which would not be the case for exoplanetary transits. However, if the depth variation was small, the light curve would resemble that of a transiting planet with a period half that of the binary star. Alternatively, if one of the events caused a small enough dip in the light curve (perhaps due to eccentricity in its orbit) it might be mistaken for the secondary eclipse of the planet.

In such cases, the distinction between grazing binary and exoplanet can usually be made by visual inspection of the light curve and identification of a V-shaped grazing eclipse or a U-shaped transit. However, exoplanets with a large inclination angle or those orbiting stars with a large amount of limb darkening (see Section 1.3.2.3) may produce light curves capable of falling into either category (depending on the quality of the data). In such a case, follow-up radial velocity measurements would reveal the masses of the bodies in the system, allowing a definitive distinction to be made.

Another potential false positive is an eclipsing binary blend; whereby the light from an eclipsing binary star[1] falls on the detector close to light from a far brighter star. In this instance, the light from the two systems will appear as one object in the resulting exposures (see Figure 1.8). Therefore, the eclipse of the binary star will only appear to reduce the overall light of the composite 'star' by a small percent of the total; as if a planet had transited the brighter of the systems.

[1] A two-star system in which one star obscures the light from its companion once each orbit - though blocking far greater light than an exoplanet.

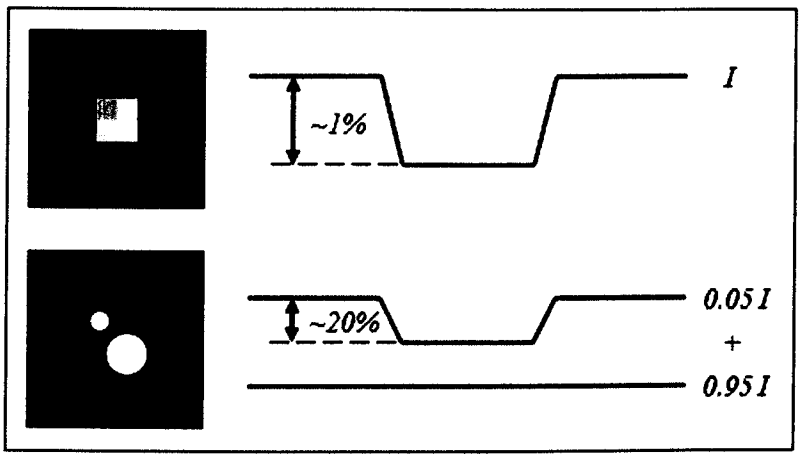


Figure 1.8: *The false detection of an exoplanetary transit caused by a binary star blend in which a low resolution image (top, left) produces a light curve with a* $\sim$ 1% *dip in intensity (top, right); characteristic of an exoplanet. However, a follow-up higher resolution image (bottom, left) reveals that the object in the low resolution image was in fact two objects - a bright object and a fainter, eclipsing binary star which exhibits a much larger (*$\sim$ 20%*) reduction in intensity during eclipse.*

Blends are most easily identified by obtaining higher-resolution images of the candidate object which would reveal the presence of both objects in the field. Alternatively, radial velocity observations of the system would reveal the co-moving spectral lines of the two stars in the binary system, together with the static lines of the brighter star.

Finally, certain types of eclipsing binary star systems can exhibit eclipses of similar depth to that of a transiting planet. White dwarf stars, in particular, can be roughly equivalent in size to gas giant planets and so would block an equivalent proportion of light during eclipse. However, such objects would themselves contribute a significant proportion of the total light emitted by the system and so would produce a far larger secondary eclipse than a planet of the same mass.

Furthermore, the mass difference between the two scenarios is more than enough to easily distinguish between them in radial velocity follow-up observations.

Quantifying the occurance rate of false detections is difficult, in part, due to the dependency on the instrumentation - the freqency of blends, for example, is partly dependent on the resolution of the telescope. However, early transit surveys typically found that 20% of potential transits were caused by planetary bodies, with the remaining 80% resulting from grazing eclipsing binaries, blended binaries and the transits of planet-sized stars.

1.3.2.5 Brown Dwarfs

Another source which, depending upon context, may or may not be deemed a false detection are brown dwarfs. A brown dwarf is an object of near-stellar mass; large enough to undergo deuterium fusion at its core, but lacking the mass to fuse hydrogen (Spiegel et al., 2010). Since the object is, nevertheless, capable of sustaining a nuclear reaction, it is technically considered a star.

However, since the exact mass-limit at which such a reaction would start is unknown, it is not always possible to state whether a detected object is a large planet or a small brown dwarf star. Indeed, the case may be that there is no precise distinction between planets and brown dwarfs and that the two classifications simply occupy arbitrarily defined segments of the same continuum of object mass.

1.3.2.6 Transit Surveys

The most successful transit survey, to date, is the Super-Wide Angle Search for Planets (or 'SuperWASP') with which this research was conducted. WASP has so far announced the discovery of, at time of writing, 48 exoplanets; accounting for over one-third of the total. WASP announced its first planets (WASP-1b and WASP-2b) in 2007 (Cameron et al., 2007) the latter of which will be discussed further as part of this research (see Chapter 7).

Alongside WASP are several other large-scale transit surveys which have helped expand the known population of transiting planets. These include the Hungarian Automated Telescope network (HAT, 30 planets), Transatlantic Exoplanet Survey (TrES, 4 planets), the XO project (5 planets) and those planets discovered as part of the larger Optical Gravitational Lensing Experiment (OGLE, 8 planets). More recently, transit surveys have moved to bespoke-built space telescopes with two currently in operation: namely Convection Rotation and planetary Transits (CoRoT, 17 planets) and Kepler (15 confirmed planets and 1235 candidate planets) (Schneider, 2010).

1.3.2.7 Drawbacks and Detection Bias

The transit method preferentially detects planets on small, short orbits or planets with large radii compared to their parent stars. As would be expected, such planets create deeper transits which are easier to detect when compared with smaller planets. However, the effects of the orbit on this detection bias is worse still.

Firstly, planets on short periods will transit their host stars much more often than those with long periods - increasing the probability that a transit will be observed. Secondly, and more significantly, planets with small semi-major axes are much more likely to transit their parent stars because such planets may still transit with a relatively small inclination angle (i). The further a planet from its star during eclipse, the closer i must be to 90° for an eclipse to occur. These factors combined have resulted in the population of planets discovered by transit being even more skewed towards the most extreme of the Hot Jupiters than those discovered by the Radial Velocity method.

1.3.2.8 Transit Timing

For a transit light curve of sufficient precision, the contact points of the transit can be determined to high precision. As a result, the exact ephemeris of the transit can be measured to the limit of the temporal-resolution used in the observations. This allows for any variations in the orbit of the planet to be assessed over multiple transits; specifically variations caused by other bodies in the system.

In any simple two-body system, comprising simply a star and one planet, both objects orbit their mutual center of mass in a steady manner. However, if a third body (such as a second planet) was introduced to the system, all three would now orbit a slightly repositioned center of mass. In a transiting system, this would cause slight variations in the timing of each transit as presence of the second, unseen, planet causes both the star and the first planet to be in slightly different positions during each transit event; determined by its own position in

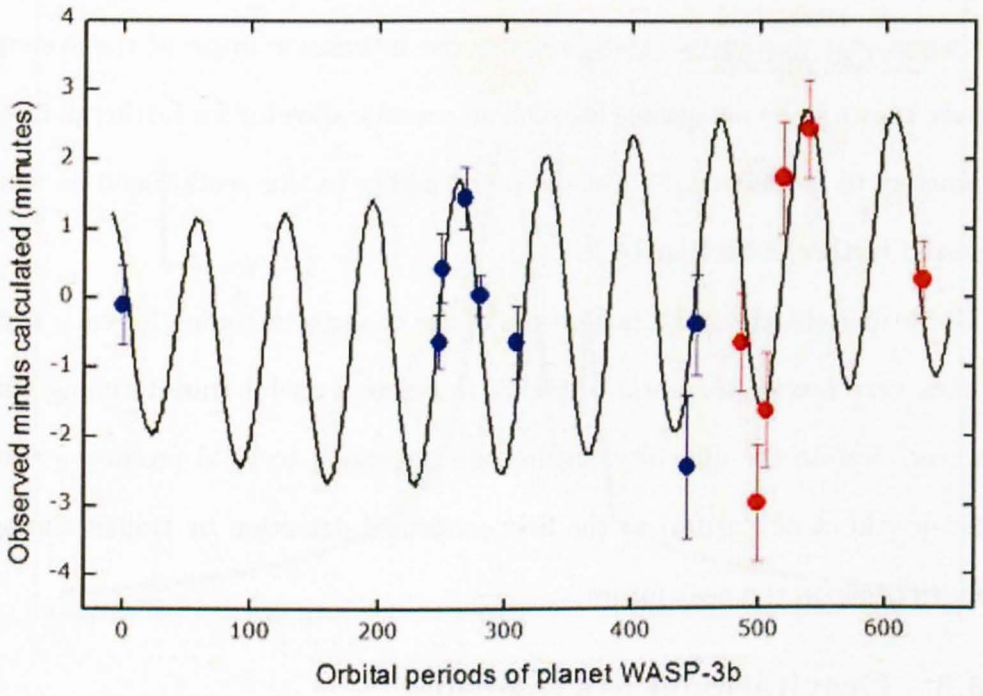


Figure 1.9: *A plot of transit timing variations (TTVs) of the exoplanet WASP-3b caused by a second planet in the observed system (WASP-3c). Before the exact period of the second planet is known, a sufficient number of points must be recorded to allow the curve to be folded with time to find the best-fit orbital period. (Maciejewski et al., 2010b)*

the system.

By measuring the timing of a succession of transits, it is possible to detect these variations and infer the orbit and mass of the second planet (Fabrycky, 2008). Plotting this variation against time (see Figure 1.9) produces a sine curve similar to that created by the Radial Velocity technique; with the amplitude proportional to the mass of the planet and the period equal to the orbital period of the new planet.

In transiting systems, it is highly likely that such a planet would also transit the parent star during its orbit (because the inclination angle of the system is already known to be favourable for such an event[1]), allowing for further planetary parameters to be derived, as was discussed earlier in this section and as will be discussed further in Section 1.4.2.

Unfortunately, the small magnitude of the changes in timing for each transit requires very precise temporal precision to create a useful transit timing curve. However, despite the difficulty, studies are beginning to yeild promising results (Maciejewski et al., 2010b) so the first confirmed detection by transit timing is likely to occur in the near future.

1.3.3 Gravitational Microlensing

Contrasting with both radial velocity and transit detection, Gravitational microlensing preferentially discovers planets far from their parent stars. A microlensing event is a short-lived increase in the intensity of a star (or other light source) caused by the gravitational field of a foreground object acting as a lens to bend more light towards the observer. Applying this phenomenon to the search for extrasolar planets involves searching for the rare occurance in which two stars cross each other in the sky (Bond et al., 2004).

In such an event, the foreground star lenses the light of the background star; causing the total light intensity of the two to temporarily appear far greater

[1]This likelihood would also be dependent on the eccentricity of the second planet and assumes that all planets in the system orbit in the same plane about their parent star - a common requirement of formation models which has, so far, been sustained by the 56 multi-planet extrasolar systems discovered to date.

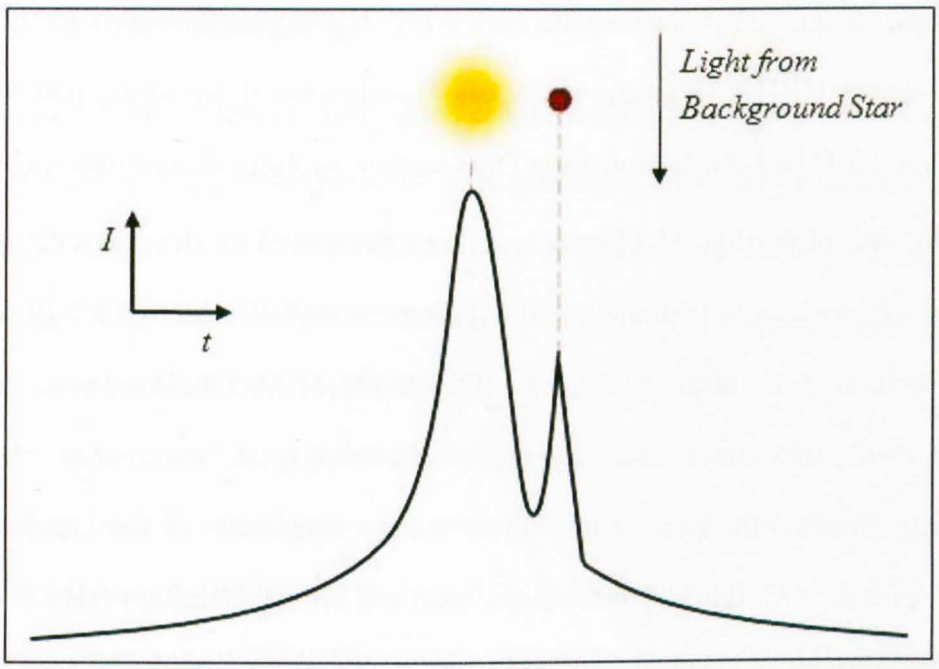


Figure 1.10: *The change in intensity during a microlensing event produced by a star and planet as they move across a background star. As the gravity of the foreground star lenses the light from the background star, the measured intensity of light increases dramatically - much more than simply the intensity of both objects combined. Then, as the planet moves in front of the background star, its gravity causes a second lensing event; producing the second, smaller spike to the right of the main peak.*

than simply the sum of the light from each. In the case where a planet is in orbit about the foreground star, a second microlensing event will occur when the planet crosses the background star. Measuring the magnitude of this peak allows for calculation of the precise mass of the planet whose gravitational field created it. The resulting change in observed intensity over time is shown in Figure 1.10.

Due to the rarity of such an event, microlensing searches adopt a similar pattern to transit searches; observing a large area of sky to capture as many events as possible. Thus far, all microlensing planets have been detected by two

surveys, often in collaboration with each other: the Optical Gravitational Lensing Experiment (OGLE, a study which has also discovered transiting planets, see Section 1.3.2) and the Microlensing Observations in Astrophysics survey (MOA).

At time of writing, 11 planets have been discovered by the gravitational microlensing technique (Schneider, 2010); between $0.66AU$ (MOA-2007-BLG-192-Lb; Bennett et al. (2008)) and $5.1AU$ (OGLE-235/MOA-53b; Bond et al. (2004)) from their parent stars. Since the degree of lensing (and, by extention, the percentage increase in light) is dependent on the magnitude of the gravitational field, planets with different semi-major axes will produce lensing events of different magnitudes, making planets with certain orbits easier to detect than others (Gould and Loeb, 1992).

While gravitational microlensing is capable of detecting a wholly different class of extrasolar planet to the far more successful radial velocity or transit methods, the technique suffers from an inability to repeat the test or to easily verify the presence of the planet. With microlensing events so rare, the chances of a repeat observation are small enough to be negligible. For microlensing planets at large orbital separations or with relatively low mass, follow-up radial velocity measurements would have difficulty detecting the resulting motion of the parent star and, in the case of large semi-major axes, might require many years of observations before a sufficient proportion of the orbit has been observed. However, in such instances, an alternate detection method may soon present a viable means of confirming such detections.

1.3.4 Direct Imaging

All previously noted detection techniques have relied upon measurements of secondary factors to infer the presence of a planet by its effect on the system it occupies (even transits look for the absence of light rather than the light from the planet itself). However, until recently, the prospect of imaging an exoplanet directly seemed more than technology was currently capable of.

On the contrary, as technology and innovation has adapted to the ever-expanding hunt for exoplanets, instruments like coronographs and techniques such as adaptive optics have been turned to the cause. Since the first confirmed discovery of exoplanet 2M1207b in 2004 (Chauvin et al., 2004), 14 exoplanets have been detected in 11 systems by direct imaging (Schneider, 2010). However, it was not until 2008 and the detection of Fomalhaut-b (Kalas et al., 2008) and the three planets in the HR8799 system[1] (Marois et al., 2008) that the orbits of any such planets were properly constrained.

Direct imaging preferentially detects planets with large semi-major axis which are easier to discern from the light of the star itself (which still dominates such images, even after the application of a coronograph to block out the star itself). However, such detections are far from fast; requiring at least two images of the candidate planet to be taken at times separated by enough to allow the planet to move along a noticable portion of its orbital path (see Figure 1.11). Without this evident movement, the 'planet' could be no more than a background star or an artifact of the detector or imaging technique being applied.

[1]Now known to harbor a fourth planet (Marois et al., 2010).

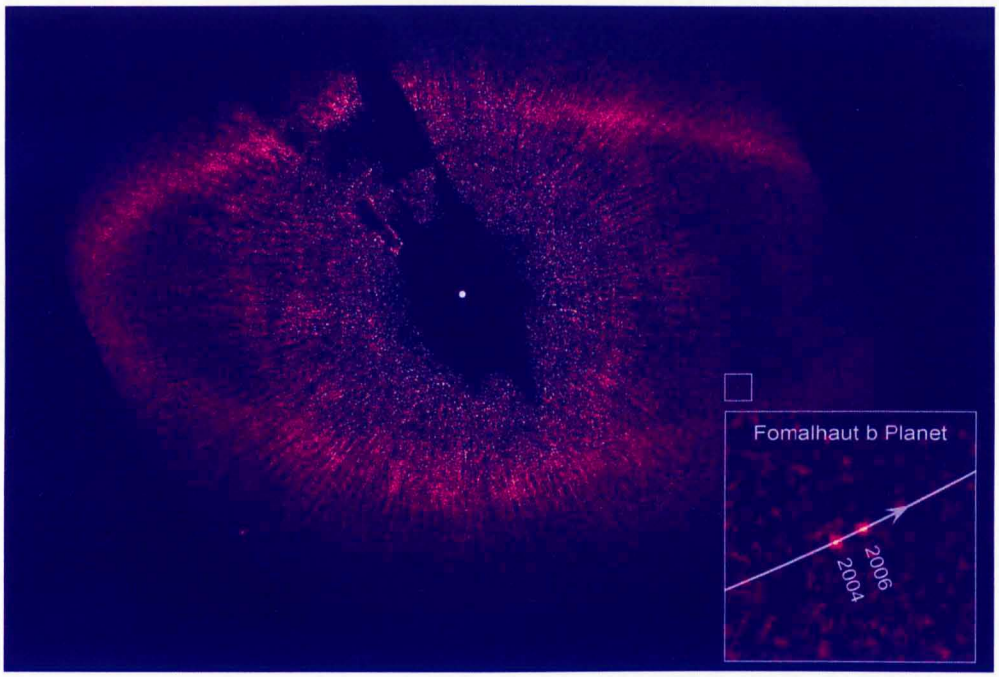


Figure 1.11: *A composite of two images showing the detection of the planet Fomalhaut b. The single images show the planet at two points in its orbit; proving that the object is indeed a planet and not a background star (which would not move with the same proper motion as the parent star) or an instrument artifact (which would be unlikely to reappear with the same intensity in two subsequent images).(Kalas et al., 2008)*

While still its infancy, technology is likely to propel the direct detection of exoplanets to the fore. In addition to instruments at the Keck, Gemini and Subaru telescopes responsible for the current population of directly imaged planets (not including HST observations performed in conjunction with ground-based studies) new instruments specifically designed for such research will soon begin taking data. The first such instrument is the Gemini Planet Imager (GPI, an addition to the aforementioned Gemini telescope in Chile) followed by SPHERE (Spectro-Polarimetric High-contrast imager for Exoplanet REsearch), an instrument for the VLT (Very Large Telescope, also in Chile). Further into the future, the E-ELT (European Extremely Large Telescope) is currently planned to host the EPICS instrument (ExoPlanet Imaging Camera and Spectrograph) which may be capable of directly imaging terrestrial-sized planets.

With every improvement in telescope power, the range of semi-major axes of detectable planets will move closer to the parent stars. Simultaneously, new technologies will emerge to allow for spetrographic analysis of such planets. However, with the number of detections only recently reaching double figures, it is still likely too soon to speculate.

1.3.5 Pulsar Timing

While not readily comparable to the aforementioned techniques, those planets detected by pulsar timing were the first known exoplanets in history. Unlike previously covered techniques, pulsar planets are not resident around solar-type stars or even stars on the main sequence.

A pulsar is a rapidly rotating neutron star with an immense magnetic field driving a set of radio jets from its magnetic poles. As the pulsar spins (with a period typically in the order of milliseconds, though pulsars have been discovered with longer periods of several seconds) these jets sweep around in very precise arcs. From the perspective of an observer in the path of this arc, the pulsar would appear to emit pulses of light as the observed jet rotates to face the observer and rotates away again.

The extreme mass and relatively minute size of neutron stars produces an extremely stable rotation; allowing for the timing of the pulses to be determined, in some cases, to nanosecond precision (Hewish, 1970). However, as with the radial velocity technique, the presence of another massive body in the pulsar system will interact with the star and perturb its rotation. The result is a periodic variation in the recorded frequency of the neutron star's pulses equal to the period of the planet. Again, as with the radial velocity technique, the mass of the perturbing planet can be calculated from the amplitude of these variations.

The first pulsar planets, around PSR1257 + 12, were detected in 1992 with masses of $0.022M_{\oplus}$, $4.132M_{\oplus}$ and $3.814M_{\oplus}$ (Wolszczan and Frail, 1992). Because of the precise timing of pulsars' rotation, the limitations on the mass of bodies which can be detected is far lower than with any other techniqe. However, with the extreme nature of the parent star, it is unlikely that these worlds are in any way comparable to any other class of extrasolar planet.

1.4 The Unanswered Questions

As the field of exoplanet hunting pushes forwards - seeking to gain an understanding of the population as a whole - unexpected discoveries have resulted in many unanswered questions. While many such questions may be answered as our knowlegde of the full population resolves, others have spawned new theories and fields of research attempting to understand and explain what has been found.

1.4.1 The Density Problem

As discussed in Section 1.3.2 it is possible to calculate both the precise mass and radius of exoplanets which transit their parent stars. Knowing both these quantities allows us to derive the mean density of the exoplanet. However, performing this calculation for the population of known transiting planets reveals an odd property of certain hot Jupiters.

Figure 1.12 demonstrates the masses of various transiting exoplanets in relation to their radii - in essence, demonstrating the range of densities covered by the current exoplanet population. The overplotted lines denote the expected densities of planets for given atmospheric make-ups with the uppermost line representing the expected radius of a planets made of pure hydrogen for a range of masses. However, from the plot it is eminently clear that several planets, most notably TrES-4b, have radii far higher than this line - indicating an astonishingly low density.

Clearly such planets cannot be made of an element with less mass than Hydrogen. Therefore, we must consider some process by which they could maintain

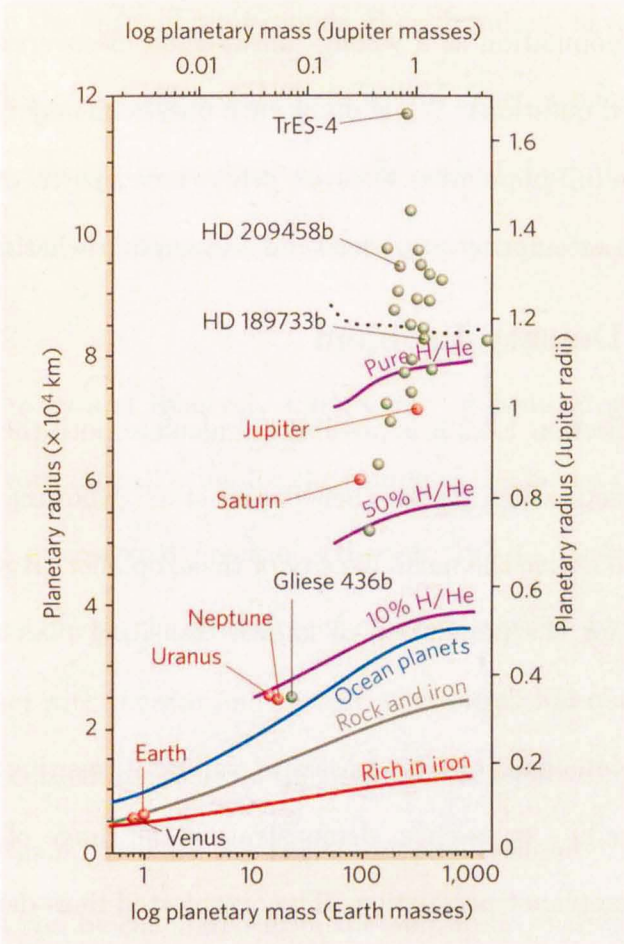


Figure 1.12: *The relative masses and radii of a selection of planets overplotted with the modelled radii for planets of different compositions. The upper, dotted line represents the radii of a planetary mass made of pure Hydrogen (though such a planet could not realistically ever exist). Many exoplanets have radii above this line - indicating that an internal process is somehow trapping heat and keeping the planets' atmospheres artificially inflated. (Sasselov, 2008)*

a temperature greater than that predicted based on the incident stellar radiation alone. It has been theorised (Fortney et al., 2008) that such planets may possess hot stratospheres with significant quantities of Titanium- and Vanadium-oxide (TiO and VO). The first tentative detection of both TiO and VO came in 2008 (Désert et al., 2008), however no more concrete evidence has yet been uncovered.

1.4.2 Atmospheric Analysis

In the case of transiting exoplanets, it can be possible to discern the make-up of the planet's atmosphere with the right instrument and observing strategy. Such research lends valuable information to studies of the aforementioned density problem, as well as aiding the fields of planet formation and dynamics.

Examining the transit event in more detail, our previous model of the transiting planet as a solid disk can be considered an oversimplification. In reality, some proportion of the light from the star will pass through the planet's atmosphere and, in the case of hot Jupiters whose atmospheres are large and diffuse, this proportion would typically be significant (see Figure 1.13).

Light passing through the atmosphere of the planet will interact with the constituent molecules therein - being scattered and absorbed dependent on their wavelengths and the species of molecules encountered. The light which escapes the far side of the planet's atmosphere is therefore imprinted with the planet's transmission spectra.

In a transit light curve, this imprinting manifests as a variation of the transit depth with wavelength. At wavelengths where the planet's atmosphere absorbs

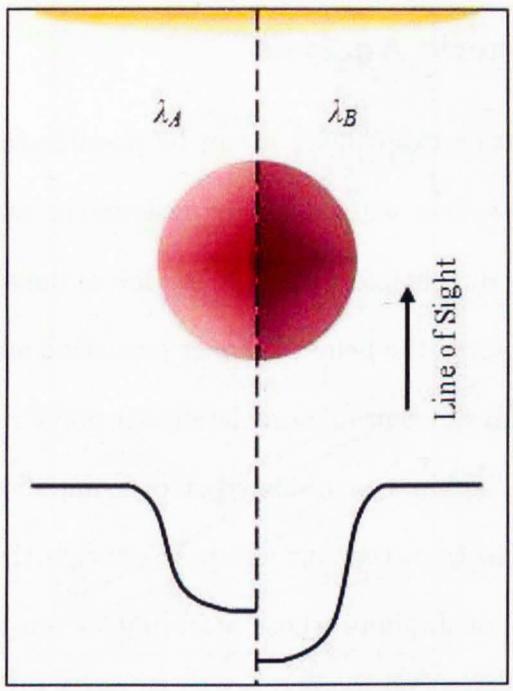


Figure 1.13: *The observed transit depth of the same planet at a wavelength where the planet is relatively transparent (λ_A, left) and relatively opaque (λ_B, right).*

incident light, fewer photons pass through the atmosphere and so more overall light is absorbed. This has the effect of deepening the transit; increasing the measured planetary radius (R_P) according to the degree of absorption. Therefore, by measuring the transit depth at a range of wavelengths it is possible to determine the transiting planet's transmission spectrum and so determine which molecules are present in its atmosphere.

To date, various elements and compounds have been detected in the atmospheres of hot Jupiters using this technique. Studies carried out by the Hubble Space Telescope (HST) and the Spitzer Space Telescope have reported the detection of water (Beaulieu et al., 2010, HD209458b), ionised hydrogen (Vidal-Madjar et al., 2003, HD209458b), sodium (Charbonneau et al., 2002, HD209458b), methane (Swain et al., 2008, HD189733b), magnesium (Fossati et al., 2010b, WASP-12b), carbon monoxide (Tinetti et al., 2010, XO-1b) and potassium (Christiansen et al., 2010, HAT-P-7). More recently, ground based observations have also proved effective in this area, confirming the earlier detection of sodium in the atmosphere of HD209458b (Snellen et al., 2008) and finding the same element in HD189733b (Redfield et al., 2008). Further discussion of these discoveries will be conducted in Chapter 2.

It is theoretically also possible to achieve a spectrum of a terrestrial-type (Earth-like) planet, though not through this same technique. Because of their much smaller atmospheres, the change in flux due to such planets' transmission spectra would be immeasurable with current technology. However, instead, their molecular properties may still be discernable by analysis of the light reflected and

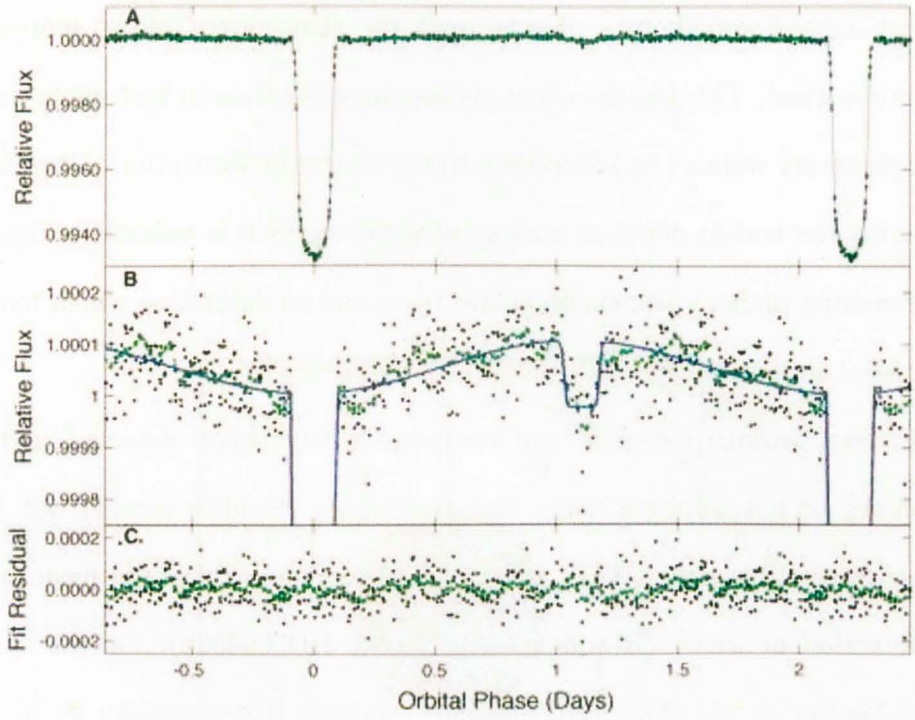


Figure 1.14: *The complete light curve of a star with a transiting planet (HAT-P-7b, observed by the Kepler space telescope) showing the primary transit at* 0.0 *of the orbital phase and the smaller secondary eclipse at* ~ 0.5. *(Borucki et al., 2010)*

emitted by the planet itself.

For most transiting planets, there are two events in each orbit; the first being the transit of the planet across the star and the second being the eclipse of the planet by the star itself. This event, known as the 'secondary eclipse' (see Figure 1.14) typically occurs one-half of the period from the transit (though, depending on the inclination and eccentricity of the system, it may occur slightly earlier or later; or, in certain conditions, it may even not occur at all).

During the secondary eclipse, the light from the planet is obscured by the star; allowing for the amount of light previously received from the planet to be

deduced from the resulting change in recorded flux. The flux from the planet itself is a combination of the reflected light from the star and the thermal flux of the planet. Different planetary conditions and observing strategy can affect which is the predominant source of planetary light. For example, observations in the infrared would be particularly sensitive to the thermal flux which would not be as significant in visible or ultra-violet wavelenghts. In contrast, planets with a high albedo would reflect more of the light from the star, reducing the impact of the thermal flux.

In either case, since the radius of such a planet would already be known from the primary transit, its geometric albedo could be approximated as:

$$A = \frac{I - I_{Eclipse}}{I - I_{Transit}} \tag{1.24}$$

where A is the planet's albedo; I is the intensity of light from the star; $I_{Transit}$ is the intensity of light during transit (where the planet's thermal emission would still be detected); and $I_{Eclipse}$ is the intensity of light during the secondary eclipse.

Hot Jupiters are expected to have a very low albedo (of order ~ 0.1), as very little light is reflected by the extremely hot, diffuse atmosphere. As a result, the planetary spectra recorded during secondary eclipse is likely to be dominated by thermal emission. To date, the secondary eclipse light curves of several Hot Jupiters have been measured (Christiansen et al. (2010); Todorov et al. (2010); Deming et al. (2010)).

In contrast, the albedos of rocky, terrestrial planets will be higher than their

hot Jupiter counterparts; resulting in a greater difference in flux during secondary eclipse when compared to the primary transit. Analysing this flux-difference at different wavelengths would reveal any aditional absorption or emission by different layers in the planet's atmosphere and, depending on the depth of said atmosphere, the surface of the planet itself. However, telescope technology has not yet advanced to the stage where such analysis is possible at the required precision and wavelength resolution (see Section 10.3).

Through such analysis, it has proved possible to discern variations in radition with longitude in the atmosphere of hot Jupiters; the result of temperature variations as heat from the incident stellar radiation is redistributed to the nightside of the planet (Knutson et al., 2007). Further discussion on the nature and detection of exoplanetary atmospheres will feature in Chapter 2.

1.4.3 Planet Formation

A more long-term hope in the search for planets is that a better understanding of the extrasolar population will aid in the development of a universally accepted theory of planet formation. At present, two leading theories exist to explain the creation of planets from the disk of gas and dusk which surrounds all new-born stars; core accretion and gravitational instability.

1.4.3.1 Core Accretion

The core accretion theory proposes that planets form from smaller bodies in a process starting at the granular level. In ideal conditions (within the correct temperature range and with appropriate density of the material in the dust disk)

the gas and dust can collect together into grains which, in turn, collect together into larger objects until those objects (at this point known as 'planetesimals') become large enough to begin accreting matter through their own gravitational pull.

At this point planets may either enter a period of runaway growth whereby they attract large volumes of gas or dust and reach enormous size or, alternatively, they may rapidly exhaust the material within their orbital range of the young star and cease growth altogether. As time progresses, the newborn star will begin to emit its solar wind of high-energy plasma and particles, which will proceed to expel any remaining dust and gas from the solar system, leaving only the formed planets (Ribas and Miralda-Escudé, 2007).

Regions of the disk too close to the parent star are typically too chaotic to allow grains of any significant size to form; both a result of their high temperature and the orbital velocity of their constituent matter. Ideally, grains require ice to grow beyond a certain size and so planet formation is expected to progress at a faster rate beyond what is known as the 'frost line'; the distance from the young star where the disk temperature falls below $0°C$.

1.4.3.2 Gravitational Instability

The gravitational instability model (sometimes simply called the disk instability model) predicts the growth of planets through regions of high density in the planet-forming disk. In this model, gravitational interactions between the young star and nearby objects (such as other stars) create density waves through the

disk, forcing material together into rings of above average density (as can be seen, on a far smaller scale, in the rings of Saturn).

Such a band within the circumstellar disk then begins to form into a planet; ejecting much of the material to account for the loss of angular momentum as the ring collects into a single point (Youdin and Shu, 2002). Observations of dust disks around young stars have revealed the existence of such rings and, the direct detection (see Section 1.3.4) of certain exoplanets (in particular Fomalhaut-b (Kalas et al., 2008)) have now revealed the presence of planets on the inside of these formations.

Chapter 2

Instrumentation

2.1 Chapter Outline

The following chapter details the instrumentation and data used in this research together with a review of comparable studies of exoplanetary transits and atmospheres. An overview of the Hubble Space Telescope's capabilities in this field, divided by instrument, puts previous studies in context. The chapter concludes with details of the data aquisition for the NICMOS instrument used in this research and a more in-depth review of previous NICMOS results.

2.2 Observing Exoplanetary Transits

As discussed in Chapter 1, observations of the transits of exoplanets can yield far more information on those planets than observations conducted through any other method currently in use. As such, transit studies typically produce the most significant new results for the field as a whole.

However, as discussed in Section 1.3.2, the rarity of transit events forces detection surveys to cover a large area in order to yield a worthwhile haul of planets.

The approach inevitably compromises on the resulting quality of the transit light curves; greatly limited the precision of the subsequently derived planetary parameters. While the majority of planets are confirmed through follow-up oberservations on superior instruments, relatively little is known about them beyond their base planetary properties (and those usually still subject to significant uncertainties). Therefore, for more detailed observations, those planets which prove most interesting often become the subjects of continual observations by better instruments.

One such instrument is the Hubble Space Telescope (STScI, 2010) which has, at time of writing, taken high-precision transit observations of eleven exoplanets: HD209458b, HD189733b, WASP-1b, WASP-2b, WASP-3b, WASP-12b, TrES-4b, XO-1b, XO-2b, GJ436b and HD149026b.

2.3 The Hubble Space Telescope

The Hubble Space Telescope (or HST, see Figure 2.1) was delivered into low-Earth orbit by the space shuttle Discovery on April 24th 1990. At 13.2m long and weighing over 11 tons (almost 10% of which is the telescope's main mirror alone), Hubble is one of the most well-known satellites in existence.

Space-based observations have traditionally exceeded those taken from the ground. Free from the absorption and interference of the Earth's atmosphere (not to mention the drawbacks of cloud-cover, light-polution and daylight), space telescopes like HST have long been capable of precision impossible from the ground. In recent years, the advent of technologies like adaptive optics have

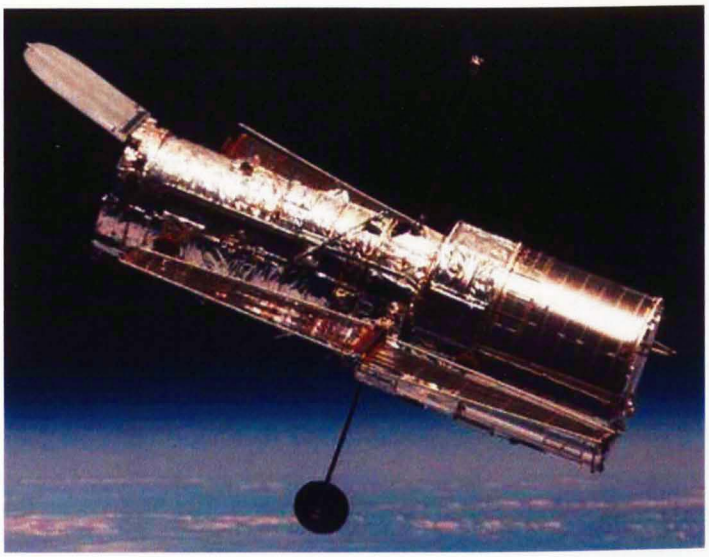


Figure 2.1: *The Hubble Space Telescope. (HubbleSite, 2010)*

allowed ground-based facilities to close the ground with their space-based rivals. However, Hubble and other such instruments still have the edge; especially in the infrared wavelength range[1] of the data covered in this thesis.

Hubble's position in low-Earth orbit means that, for a significant period of each orbit, the telescope will be on the wrong side of the Earth to observe its target. Rather than slewing to a new position for this time (which would reduce the pointing accuracy of the telescope and therefore the quality of both sets of observations) Hubble simply stays dormant. As a result, each set of Hubble observations can be split into orbits, each separated by a gap while the telescope waits for the Earth to move out of its way. One full orbit lasts 96 minutes, meaning that 48 minutes can typically be spent observing the target.

[1] Atmospheric water absorbs much of the infrared light incident on the Earth, greatly limiting the precision of ground-based observations at these wavelengths.

Unfortunately, the typical duration of an exoplanetary transit roughly 2-3 hours; meaning that Hubble can never capture a complete transit event. Therefore, in order to acheive a complete transit light curve, Hubble must execute multiple observations of the same target (also known as 'visits' of observation to that target) scheduled so that different orbits of HST capture different portions of the complete transit light curve. These observations can then be folded together by the planet's orbital period and combined to form one complete light curve.

2.3.1 STIS

At the dawn of the era of transiting planets, many researchers saw Hubble as the perfect instrument for transit follow-up studies. The most applicable instrument onboard HST was the Space Telescope Instrument Spectrograph (STIS) which replaced the Faint Object Spectrograph (FOS) and High Resolution Spectrograph (HRS) as part of the second HST reservicing mission in 2004 (Proffitt and et al., 2010).

STIS made a number of early exoplanet breakthroughs - including the first detection of an exoplanetary atmosphere; that of HD 209458 b (Vidal-Madjar et al., 2003). Unfortuntely, just as the full potential of the instrument was being realised, STIS suffered a power supply failure in 2004 and went offline. The instrument was restored during the final HST servicing mission in 2009.

2.3.2 ACS

Another HST instrument applicable to transit observations was the Advanced Camera for Surveys (ACS; Maybhate and et al. (2009)). ACS was installed as part

of the 2002 Hubble reservicing mission; replacing the Faint Object Camera (FOC; Nota and et al. (1996)) and effectively making the Wide-Field and Planetary Camera 2 (WFPC2; McMaster et al. (2008)) obsolete.

ACS possessed three channels for different purposes of astronomy: a Wide-Field Channel (WFC) designed for general-purpose, high precision observations; a High-Precision Channel (HPC) with a smaller field of view than the WFC but equiped with various photon masks and coronographs for observations of dust disks or other objects where nearby intense light sources would otherwise interfere with results; and a Solar Blind Channel (SBC) designed for conducting high-precision photometry in the ultra-violet.

ACS was designed to be a highly versatile instrument capable of observing objects across a variety of wavelengths (from near-infrared to ultra-violet) and depths of field. ACS's collection of wavelength-specific filters, coronagraphs and grisms (see Section 2.4.2) gives it the ability to perform all manner of observations on all manner of objects. ACS's Wide Field Channel in particular was the perfect tool for transit photometry.

Unfortunately, ACS suffered a number of failures; starting in the summer of 2006 and culminating in early 2007 with the failure of all three channels. Since then, both the WFC and SBC have been brought back online, but not soon enough to be applicable to early transit observations.

2.3.3 NICMOS

With ACS offline, researchers turned to an alternate Hubble instrument; the Near-Infrared Camera and Multi-Object Spectrometer (or NICMOS; Viana et al. (2009)). NICMOS was originally installed on HST during the 1997 servicing mission along side the Space Telescope Imaging Spectrograph (STIS); replacing the Goddard High-Resolution Spectrograph (GHRS; Soderblom et al. (1995)).

NICMOS operated from 1997 through to 1999 when its dewar depleted its coolant reserves. In 2002, along with the installation of ACS, NICMOS was fitted with a crycooler to replace the need for coolant. The instrument then operated through to 2008 (taking the vast majority of HST exoplanet observations in the process, see Section 2.5) until the failure of its cryocooler. Since that time, the cryocooler has been restarted, only to fail again, on several occassions. At time of writing, NICMOS is no longer availble for future HST proposals and, with its capabilities superceded by the new Wide-Field Camera 3 (WFC3), it seems unlikely that NICMOS will be rturned to operations.

The research of this thesis will deal solely with that data taken by NICMOS of the exoplanets WASP-2b, HD189733b and G436b. However, the data reduction and analysis techniques developed for this purpose (see Chapters 3 through 6) are applicable to any data set with a similar observing strategy (see Section 2.4).

2.3.4 Next-Generation Instruments

In 2009, the final Hubble reservicing mission installed two new instruments to the space telescope; replacing those instruments installed by the very first servicing

mission in 1993.

The Wide-Field and Planetary Camera 2 (WFPC2) was replaced by Wide-Field Camera 3 (WFC3, Dressel et al. (2010)); taking over the versatility of the older wide-field camera. With a collection of filters, grisms and prisms, coupled with a greater range of wavelengths and superior resolution, WFC3 effectively replaces both WFPC2 and NICMOS. At time of writing, only one transit observation has been performed by WFC3; that of WASP-3b.

The Cosmic Origins Spectrograph (COS, Dixon and et al. (2010)) replaced the Corrective Optics Space Telescope Axial Replacement (COSTAR; installed in the very first reservicing mission to compensate for the incorrect curvature of Hubble primary mirror). COS is an ultra-violet spectrograph with a sensitivity great enough to overshadown both its predecessor and the extant STIS instrument on Hubble (though STIS maintains a superior wavelength range).

In 2010, COS observed the transit of WASP-12b Fossati et al. (2010b); Fossati et al. (2010a)), acheiving the first evidence that the atmospheres hot Jupiters can be boiled into space by the incident radiation from their parent stars.

2.4 Observing Strategy

Since the NICMOS instrument was not designed for use at the precision required for useful transit photometry, care must be taken when outlining the observations in order to make best use of the instruments advantages while limiting the impact of its drawbacks.

One such advantage of NICMOS is its `MULTIACCUM` mode. The NICMOS

detector is a 256 × 256 pixel CCD (Charge-Couple Device) consisting of four 128 × 128 chips which are each read-out independently. In `MULTIACCUM` mode, NICMOS reads out its CCD several times over the course of a single exposure. This allows for the output of each pixel to be considered in terms of the rate of flux falling on it rather than simply the total flux accrued during the whole exposure. As a result, any anomalous occurances (such as cosmic-ray impacts) can be easily identified and removed from processing (Viana et al., 2009).

2.4.1 The Duty Cycle

As with any CCD constructed before the use of Charge Transfer devices[1] became commonplace, NICMOS must take some time after each exposure to read-out the array and reset the CCD ready for the next image to be taken. As a result, every exposure has a 'duty cycle' expressing the fraction of operational time is spent actively taking observations; calculated as the ratio of the exposure time to the total time between exposures (including the read-out time, the reset time and other such delays). Stated mathematically, the duty cycle is:

$$D = \frac{t_{exposure}}{t} \tag{2.1}$$

where D is the duty cycle; t is the time between exposures; and $t_{exposure}$ is the exposure time.

[1]Charge-Transfer CCDs have equal areas of exposed and shielded pixels, allowing the chip to quickly transfer the pixel values from the exposed pixels to the shielded pixels at the end of each exposure. This allows such detectors to read out the values in the shielded pixels while the instrument is already conducting the next exposure; greatly improving the duty-cycle of such instruments.

For high-precision transit observations, we require as much light as possible to acheive the best transit light curve. However, under normal circumstances and given the typical intensity of an exoplanet host star, a single image would over-expose the CCD in a matter of seconds. This limit on the exposure time would normally impose a terrible duty cycle ($D \ll 1$) where the telescope is spending the majority of its time reading out the array rather than taking the actual data. In this set-up, more than half of each visit (and therefore more than half of the available light) is being wasted. It is therefore imperative to find some other means of improving the duty cycle.

The simplest means of acheiving this goal is to prevent the over-exposure of the CCD; thereby removing the limit on exposure time. Since we are primarily concerned only with the light from the target star, we do not need to worry about the sharpness of the image itself. As a result, defocussing the camera has the effect of spreading the incident light over more pixels; thereby allowing for a longer exposure time without adversely affecting results. Defocussing the NICMOS camera can improve the duty cycle by over a factor of 10, however, there are still means to improve it further.

2.4.2 Grisms

A grism is a grated prism designed to split incident light into a specific wavelength range. Their application produces a distinct, limited-range spectrum in place of the normal image of each object in the telescope's field (hence the term 'Multi-Object Spectrometer' in the name NICMOS).

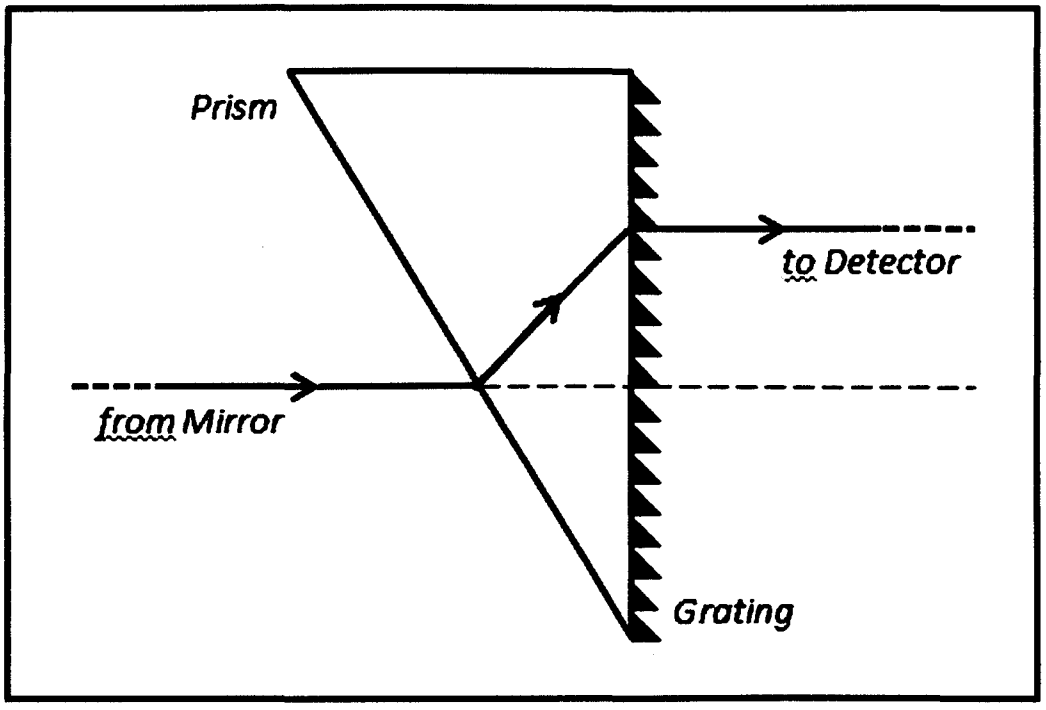


Figure 2.2: *The path of a ray of light of a given wavelength passing through a grism.*

Application of a diffraction grating alone would produce a deviation in the path of the incident light dependent on the grating in question and the wavelength of light incident upon it. To account for this, grisms mount the defraction gratings on the back of a prism designed to have the opposite effect on the light path to that of the grating (as shown in Figure 2.2. In effect, the prism corrects the light path for the wavelength range of its grating, requiring no further optics to return the observed light to the detector. This allows the grisms in NICMOS to be mounted on a wheel which can be easily rotated into the light path to conduct grism observations without requiring a separate instrument.

NICMOS has three grisms, each named for their central wavelength, covering

ranges of 0.8-1.2μm (G096), 1.1-1.9μm (G141) and 1.4-2.5μm (G206). Applying one of these grisms to an already defocussed image spreads the incident light out still further; covering hundreds of pixels and allowing for far longer exposures to be performed (Gilliland, 2005). Furthermore, since grisms spread the light by wavelength, it potentially allows for spectrscopic analysis of exoplanetary transits.

2.5 Previous Studies

Previous studies of exoplanetary transits using NICMOS have produced a variety of results; first among which came in 2008 with the announcement of the detection of methane in the atmosphere of HD189733b (Swain et al., 2008) from the G206 grism. However, the atmospheric transmission spectrum from which this result was derived appears potentially inaccurate; with grossly underestimated error bars, little account taken for systematic effects, clear artifacts of the grism's wavelength sensitivity and a key feature in the spectrum falling precisely on the chip boundary of the NIC3 camera where the bias voltage is known to be erratic. Further discussion of this result will take place in Chapter 8.

Following the 2008 result, analysis of the HD189733b secondary eclipse (from a second visit in the same data set) was published in 2009 (Swain et al., 2009c) announcing the detection of water, carbon monoxide and carbon dioxide in the dayside spectra of the planet. However, as with the claimed detection of methane, the spectrum from this result suffers from grossly underestimated uncertainties of order 10^{-5} derived from light curves with an RMS-spread almost 100 times greater.

Using the same technique on the secondary transit of HD209458b (Swain et al., 2009a), produced reported detections of methane, carbon dioxide and water. Furthermore, analysis of the XO-1b data set (Tinetti et al., 2010) reportedly also detected water, methane, carbon monoxide and carbon dioxide.

However, while one research group have apparently achieved great success in the analysis of HST NICMOS data, several other groups have, so far, been unable to duplicate their results or acheive similar success with other NICMOS grism transit observations. Research conducted by both Pont et al. (2009) and Carter et al. (2009) proved inconclusive in determining the transimission spectra of the Hot-Neptunes GJ436b (see Chapter 9) and HD149026b respectively.

More recently an extensive analysis of the HD189733, XO-1 and GJ436 data sets (Gibson et al., 2011) concluded that no definitive results could be derived from the data which they found to be still heavily influenced by instrument systematics. Further discussion on the exact analysis applied to each data set will be covered in-situ with their analysis as part of this research in the relevant Results chapters (7 - 9).

Chapter 3

Methods 1 - Data Processing

3.1 Chapter Outline

The following chapter details the initial processes of data reduction required to derive a raw transit light curve from a NICMOS NIC3 Grism data set. An overview of the whole pipeline is followed by a description of the images themselves and the programming steps required to process them - bad-pixel identification and removal; wavelength mapping; flat fielding; and background subtraction. The chapter concludes with the processing steps which locate the first-order spectrum in each data set and gather the light therein - thereby extracting a raw transit light curve.

3.2 The Data Processing Pipeline

As previously addressed in Chapter 2, HST observation data is pre-processed by the HST 'Calnic' pipeline and supplied in the standard '`FITS`' format. While this pipeline produces a variety of outputs, we are primarily concerned with the 'cal' (calibrated) data files - containing one final image, processed from all

MULTIACCUM exposures, together with pixel maps of both uncertainty estimates and data quality flags.

A typical HST NICMOS observation of an exoplanetary transit will involve two types of calibrated images - those taken through a wavelength-specific filter and those taken with a grism. Ultimately, the transit light-curve will be extracted from the grism images. However, both types of exposure are required to fully process the observations and extract a precise and believable light curve.

To robustly perform the data-processing steps required to produce such a light curve (as detailed in Chapters 3 - 6), a bespoke IDL (Interactive Data Language) pipeline has been created. This approach ensures that every exposure in each data set is processed independently and identically - thereby producing robust and readily comparable results.

The actions performed by this pipeline can be considered as four 'phases' of processing - Initial Data Reduction (detailed in this chapter), Systematic Noise Reduction (detailed in Chapter 4), Light Curve Modeling (detailed in Chapter 5) and Wavelength Dependence Analysis (detailed in Chapter 6).

The behavior of this pipeline is controlled by a '`Config`' (configuration) file unique to each data set. This file carries information such as the number of HST visits, orbits and exposures in each data set; as well as existing parameters and uncertainty estimates for the stellar and planetary properties. Each config file also contains program switches which activate and control certain processing steps. How such steps affect the actions of the pipeline will be addressed in the following chapters as each programming step is described and explained.

3.3 Initial Data Reduction

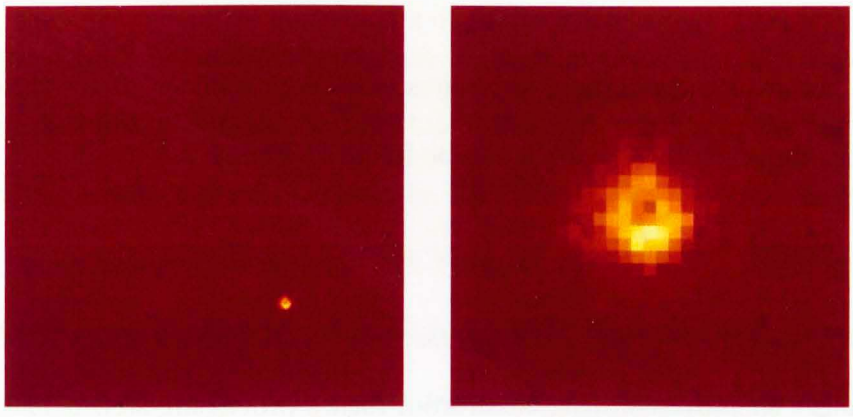
A standard HST visit of observation to a transiting exoplanet typically produces at least several-hundred grism exposures and up to a dozen direct (filter) exposures. Examples of these images can be found in Figure 3.1.

3.3.1 Bad Pixels

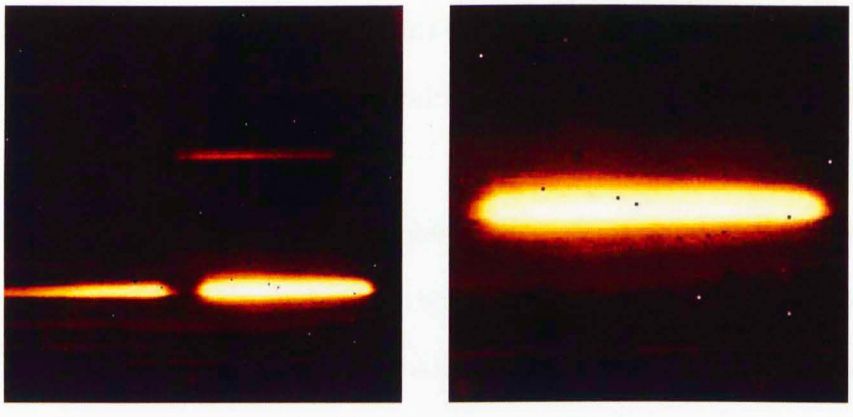
Initially, both types of image are processed to identify and flag any anomalous or bad pixels - thereby removing them from future processing steps. The majority of such pixels are easily identifiable from the Data Quality pixel map extension in each calibrated `FITS` file. However, this pixel-map is by no means exhaustive, requiring the application of further methods if we are to ensure that no bad pixels remain.

In a simple 'first-pass' over each exposure, the pipeline identifies and flags pixels which can be considered clearly anomalous. In this step, the pipeline identifies unphysical pixel values (those which are unreal or below zero), while also checking for values which are so large as to be unbelieveable. Such values are defined as those which exceed the expected flux of the target star by more than a factor of ten. This expected flux is easily calculated from the known stellar parameters and the image's exposure time. Alternatively, this threshold value can be set in the `Config` file, should the user wish to estimate it by some other means.

While this first-pass is able to identify a large percentage of anomalous pixels, it is inevitable that some such values will fall within the believable range and



(a) Direct Image



(b) Grism Image

Figure 3.1: *Two HST NICMOS NIC3 exposures. (a) A direct image taken with the F166N (1.66μm, narrow) filter shown as the full 256 × 256 image (left) and a close-up on the target star (right). Note the annular shape of the point spread function (PSF) - the result of defocussing the NIC3 optics to spread the light over a larger number of pixels (See Chapter 2). (b) A grism image taken with the G141 (centred on 1.41μm) grism shown as the full 256 image (left) and a close-up on the first order spectrum (FOS, right). The spectrum in the upper half of the image is that of a background star, too faint to be seen in the accompanying direct image.*

thus be ignored. Therefore, a second more stringent identification step must be employed.

Since we are only concerned with long-term variations in flux (namely the planetary transit, the ingress and egress of which are unlikely to be less than 15 minutes), short term variations (< 5 minutes, or those present in one image alone) - such as cosmic ray impacts and similar anomalies - would serve only to aversely affect the results. Such short-term variations can be easily identified by performing a time-line analysis on each individual pixel - effectively studying the value of each individual pixel across a complete visit of observation. In this way, we can identify and remove any pixel values which deviate by more than five standard deviations from the mean number of counts reported by each pixel over the course of a single visit. These removed values are later replaced with values interpolated from the surrounding pixels (see Section 3.3.2).

To ensure that no long-term trends are inadvertently removed in this process, we consider each point relative to a local-mean calculated from a best-fit line:

$$C_{x,y}(t_i) \simeq J_{x,y}(t_i) \quad = \quad m(t_i - t_0) + c \tag{3.1}$$

where $C_{x,y}(t_i)$ is the counts recorded by a pixel with detector coordinates (x, y) at the time, t_i, of exposure i; $J_{x,y}(t_i)$ is a linear, best-fit model to that data; m is a gradient coefficient denoting any linear-trend in the counts reported by this pixel; c the intercept of this linear trend at $(t_i - t_0) = 0$; and t_0 is the time of the

first exposure in the current visit.

Taking the standard deviation (σ) of points about this line, we can define a limit outside of which any pixel value can be considered anomalous:

$$T(t_i) = m(t_i - t_0) + c \pm 5\sigma \tag{3.2}$$

Or alternatively:

$$T(t_i) = m(t - t_0) + c \pm 5\sqrt{\frac{1}{N}\sum_{i=1}^{N}(C_{x,y}(t_i) - J_{x,y}(t_i))^2} \tag{3.3}$$

where T is the good-pixel threshold for the pixel with coordinates (x, y) at time t_i of exposure i; $J_{x,y}(t_i)$ is our best-fit line (as defined in Equation 3.3.1); m and c are the coefficients of that line; σ is the standard deviation of the pixel values about that line; N is the total number of images; $C_{x,y}$ is the counts recorded by the pixel; and t_0 is the time of the first exposure in the current visit.

Comparing each pixel with this $\pm 5\sigma$ threshold allows us to more stringently identify and remove anomalous pixel values from future calculations. A typical pixel-timeline can be found in Figure 3.2. The number of standard deviations threshold can be altered using the `Config` file.

Once identified, bad pixels can either be simply removed from processing or interpolated over to estimate a replacement value. This option can also be controlled via the `Config` file.

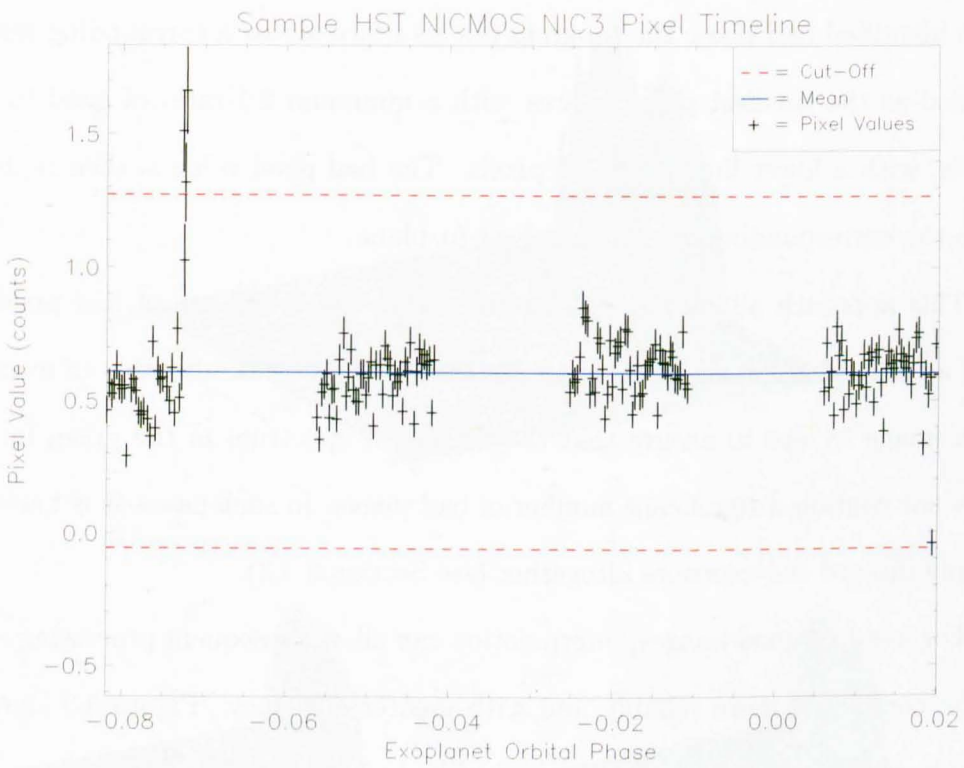


Figure 3.2: *The value of a pixel in the center of the first order spectrum over a complete HST visit of observation. The red dashed-lines show the cut-off threshold - calculated as four standard deviations from the mean (accounting for trends in the data). Any pixels outside this range are flagged as anomalous and removed from future calculations.*

3.3.2 Interpolation

Should the user choose to interpolate over the identified bad pixels, the pipeline addresses each exposure independently and uses only those pixels which it has not flagged as anomalous to arrive at estimated values for the missing data. For each identified bad pixel, the program fits a surface across a surrounding region, defined as the smallest possible area with a minimum 2:1 ratio of good to bad pixels, with a lower limit at 5×5 pixels. The bad pixel value is then replaced with the corresponding value of the best-fit plane.

This approach allows the pipeline to handle whole regions of bad pixels as well as individual anomalies. As an additional check, it is advisable to examine each image by eye to ensure that the first-order spectrum in the grism images does not contain a significant number of bad pixels. In such cases, it is better to simply discard the exposure altogether (see Section 3.3.3).

For less-maligned images, interpolation can allow subsequent processing steps to be completed more reliably and with greater efficiency. Figure 3.3 shows a sample grism image at the three stages of bad-pixel processing.

To ensure that the interpolation procedure could not potentially corrupt real data, the process was supplied with images where good pixels had been deliberately flagged as anomalous. In this instance, the interpolation procedure calculated replacement values for such pixels within three uncertainty estimates ($< 3\sigma$) of the pixels' original values. The only exception occured in cases where such pixels were grouped in regions larger than 3×3 pixels. As a precaution, images with such groups of bad pixels near the first order spectrum are automatically removed

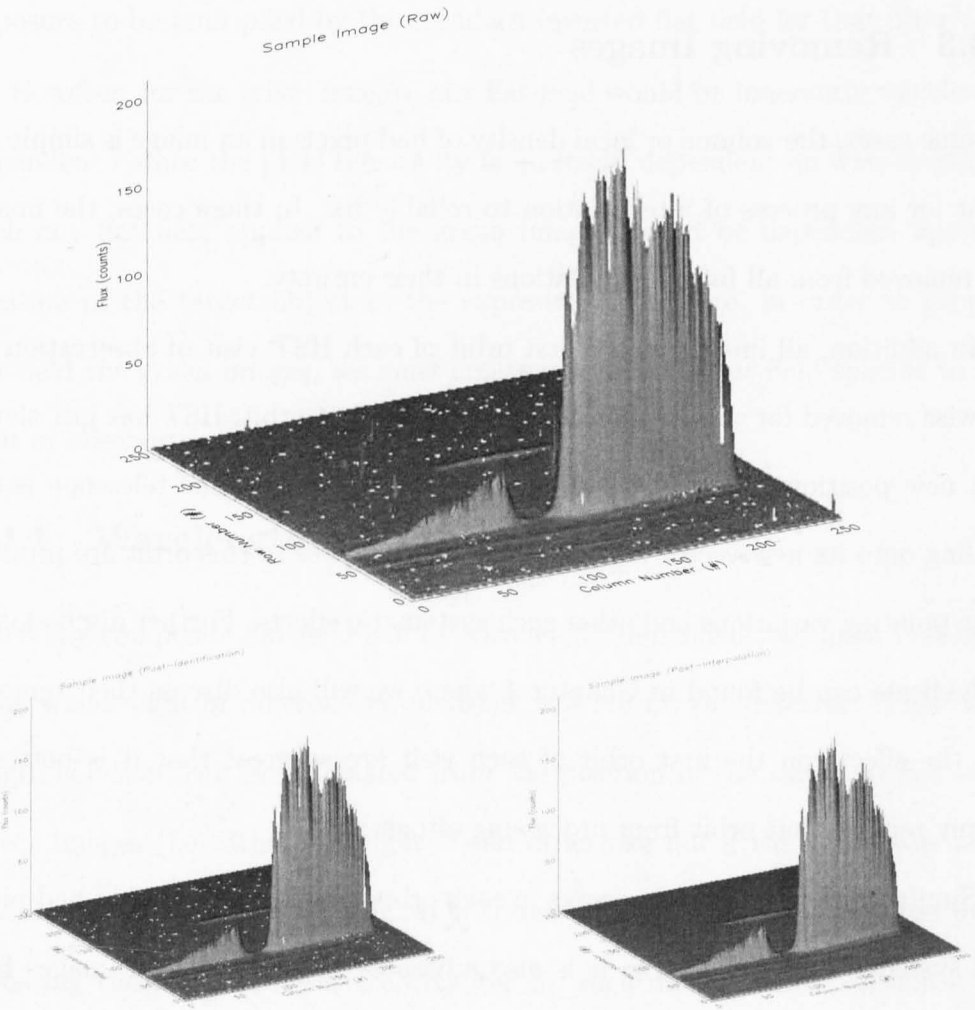


Figure 3.3: *Three surface plots showing the same image during the process of bad pixel identification and removal. The first plot shows a sample image before its bad pixels are identified. The second plot shows that same image with its bad-pixels identified and removed. Finally, the third plot shows the image after a process of interpolation.*

from processing altogether. However, for the data discussed in this thesis, no such image exists outside the first orbit of each visit of observation.

3.3.3 Removing Images

In some cases, the volume or local density of bad pixels in an image is simply too great for any process of interpolation to reliably fix. In these cases, the images are removed from all future calculations in their entirety.

In addition, all images in the first orbit of each HST visit of observation are likewise removed for similar reasons. During this first orbit, HST has just slewed to a new position to locate its target object. As a result, the telescope is still settling onto its new vector and, as such, all exposures in this orbit are prone to large pointing variations and other such systematic effects. Further discussion on such effects can be found in Chapter 4 where we will also discuss their removal, but the effects on the first orbit of each visit are so great that it is better to simply remove that orbit from processing altogether.

Similarly, the first few exposures in each orbit are likewise prone to bad pixels and systematic errors, and so it is also advisable to remove these images from processing. Typically, the first three images in each orbit are removed in this way, but the user may choose to change this number via the config file.

3.4 Flat Fielding

Once any anomalous pixel values have been identified and accounted for (whether by removal or interpolation), each exposure must then be flat-fielded to remove

any artifacts of pixel sensitivity, optical path or other such systematic effects (see Chapter 2). For the direct images in each data set, this simply requires each exposure to be multiplied by the standard inverted flat field for that filter[1].

However, for the grism images, any flat-field would be inherently wavelength-dependent - since the pixel sensitivity is, in itself, dependent on wavelength. As such any flat field applied to the grism images would be dependent upon the location of the target object in the exposure. Therefore, in order to properly flat-field the grism images, we must create a composite flat-field specific to each visit of observation.

3.4.1 Wavelength Mapping

Since any composite flat-field will be wavelength dependent, we must first determine where light of different wavelengths will fall on the detector. This 'wavelength solution' can be calculated from the position of the target object in the direct images (i.e. where all light would otherwise fall if the grism was not in place). However, since HST NICMOS transit observations are conducted by defocussing the NIC3 camera (see Chapter 2), such direct images contain a very diffuse target object with no clear center.

Therefore, in order to measure the location of the target star, we first sum together all direct images to create a single, composite direct image with a high signal-to-noise ratio (S/N). From this image, we can initially calculate a rough

[1]Flat-fields for all Hubble intruments were retrieved from the HST online resource (STScI, 2010). All such field-fields are inverted (i.e. each pixel in the flat field 'exposure' is given in units of $count^{-1}$) and, as such, each exposure is multiplied by the flat field instead of divided.

estimate of the stellar position simply by identifying the pixel with the maximum flux. This position can then be further refined by examining the column- and row-totals of this high-S/N image.

If we define the spread of the defocussed light across the detector as an integral of column- or row-totals, we can simplify that integral to a sum, such that:

$$L = \sum_{\alpha}^{\omega} C(x) \tag{3.4}$$

where L is the light from the star; $C(x)$ is the counts reported in each column, x; and α and ω are our column-boundaries on either side of the defocused star.

Then we can define the stars actual location as the 'flux-center' of the light from the star, such that:

$$\sum_{\alpha}^{S_x} C(x) = \sum_{S_x}^{\omega} C(x) \tag{3.5}$$

where S_x is the location of the star in the x axis.

In effect, we wish to solve the above equation to find the point where the light falling on both sides is equal. However, since we are dealing with discrete pixels, this approach can only tell us in which pixel the centre of the star lies. In order to arrive at a more precise figure, we apply partial pixel arithmetic to calculate the exact location of the star as:

$$S_x = P(x) - 1 + \frac{\left(0.5\sum_{\alpha}^{\omega}\right)C(x) - \sum_{\alpha}^{P(x)-1} C(x)}{\sum_{\alpha}^{P(x)} C(x) - \sum_{\alpha}^{P(x)-1} C(x)} \tag{3.6}$$

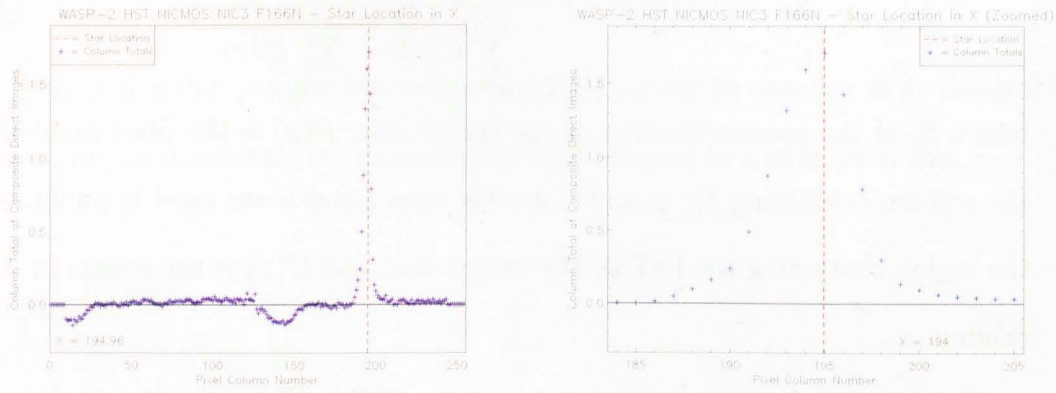
where S_x is the precise location of the target star; $P(x)$ is the pixel number of the column containing S_x; α and ω are the upper and lower pixel boundaries of the region containing the PSF of the target star; and $C(x)$ is the counts in each column, x.

Once the pixel containing the centre of the PSF has been located, the precise location of the centre can be calculated by taking the ratio of the shortfall in counts between that pixel's lower boundary and 50% of the total light from the star over the total light falling in that pixel. Performing this in both axes gives the precise location of the star, as demonstrated in Figure 3.4.

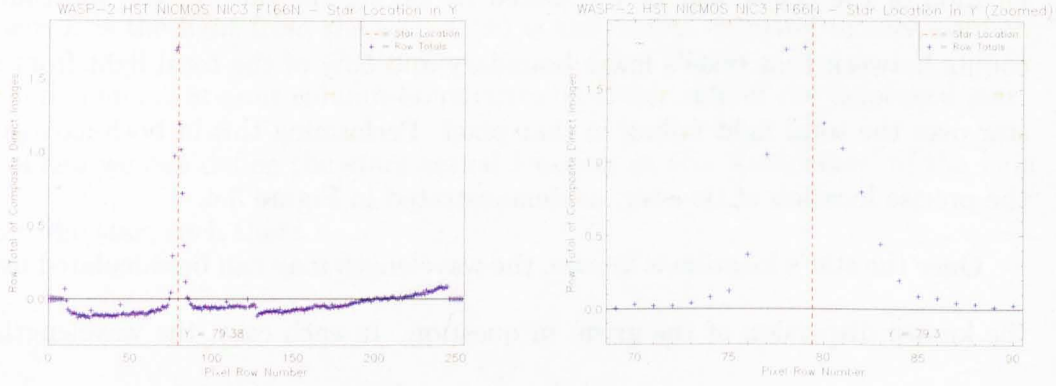
Once the star's location is known, the wavelength map can be calculated using the known dispersion of the grism in question. In each case, the wavelength of light falling in a given pixel has a linear relationship, such that:

$$\lambda = m(x - x_0) + c \tag{3.7}$$

where λ is the wavelength of light; x is the column-number (assuming the first order spectrum has been oriented to be parallel with the x-axis); x_0 is the location of the star in the direct image; and m and c are constants specific to the applied grism.



(a) x-axis



(b) y-axis

Figure 3.4: *The Column- (top) and Row-Totals (bottom) of a composite direct image (blue) used to precisely estimate the position of the defocussed target star. The plots on the left show the full width of the detector while the plots on the right have been zoomed-in to show the area immediately around the star. The vertical, dashed red line in each image demonstrates the estimated center of the star's PSF (Point Spread Function) in each axis.*

The values of m and c are specific to each grism and are provided in the HST Data Processing Handbook (Thatte et al. (2009)). Naturally, since the grism exposures are highly defocussed, each wavelength of light will be spread out across a region equivalent to the PSF in the direct images. However, since very precise values of m and c are available, this approach can at least assure a reliable measure of the central position of any given wavelength's distribution.

3.4.2 Generating the Composite Flat Field

With a wavelength map for the dispersion in the grism images, a composite flat field can be created from a selection of normalised, wavelength-specific flat-fields taken for the various NICMOS filters. In each case, these wavelength-specific flats were chosen to comprehensively cover the wavelength range of the grism in question.

The composite flat field is compiled by analysing each pixel in turn, with respect to the corresponding pixel in each wavelength-specific flat. For each pixel, the values of the wavelength-specifc flats approximate a second-order polynomial relationship with wavelength. The best value for the composite flat can therefore be found by fitting a quadratic to these values and solving this quadratic for the wavelength of light expected to fall in that column (see Figure 3.5), such that:

$$D(x,y) = a_{x,y}\lambda_x^2 + b_{x,y}\lambda_x + c_{x,y} \tag{3.8}$$

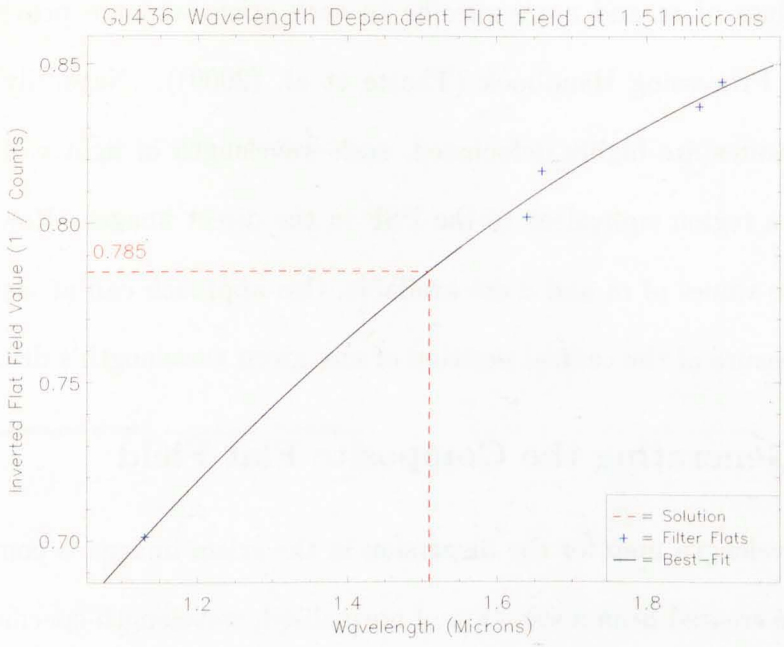
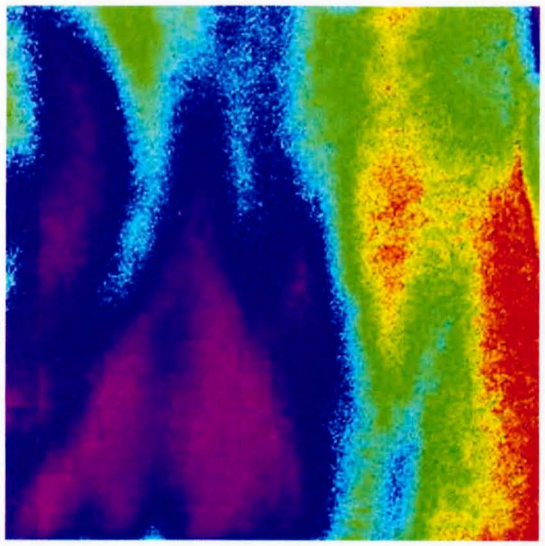


Figure 3.5: *The quadratic fit (blue) to a series of wavelength-specific flat field values (black) for a single pixel. The red line shows the solution of this quadratic for the wavelength of light expected to fall in this pixel.*

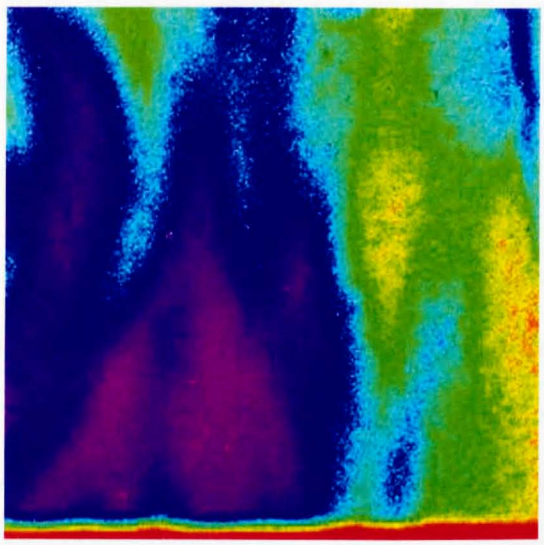
where D is the value of our composite flat field at the pixel with coordinates (x, y); λ_x is the wavelength expected to fall in the column, x; and $a_{x,y}$, $b_{x,y}$ and $c_{x,y}$ are the coefficients to the quadratic fit through the wavelength-specific filter flat fields for this same pixel.

Performing this exercise for each pixel in the NIC3 256 $\times$ 256 array produces a composite flat field (see Figure 3.6) which can then be confidently applied to each grism image in the corresponding data set.

Previous studies (Gilliland (2005)) have applied this composite flat-field only to the region immediately around the first order spectrum and have implemented a simple wavelength-specific flat (typically one from the middle of the current



(a) G141 Composite Flat Field



(b) F200N Flat Field

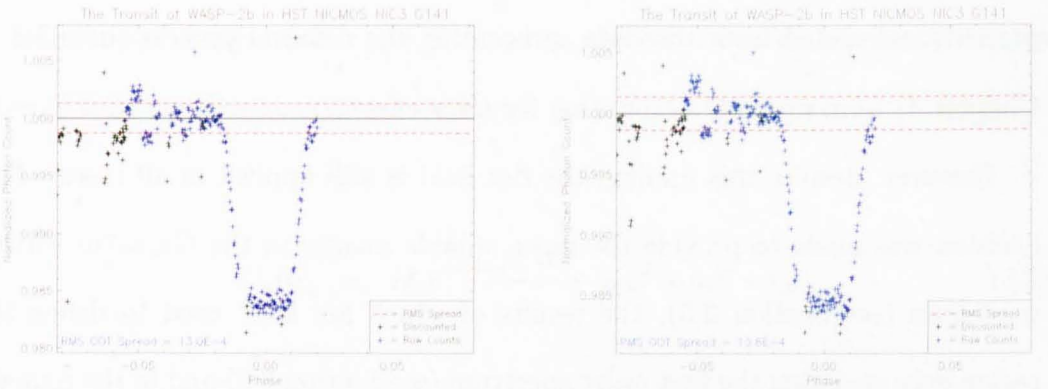
Figure 3.6: *(a) A typical composite flat field created for the G141 grism shown in contrast with (b) a wavelength-specific flat field taken with the F200N filter.*

grism's wavelength range) to flat-field the rest of the image. However, given the importance of accurately locating the first order spectrum (see Section 3.6) and the fact that any definition of the first order spectrum's 'local region' at this stage would be purely arbitrary, the decision was made to forgo this splicing of the flat fields and to simply implement the composite flat field to the entirety of each image.

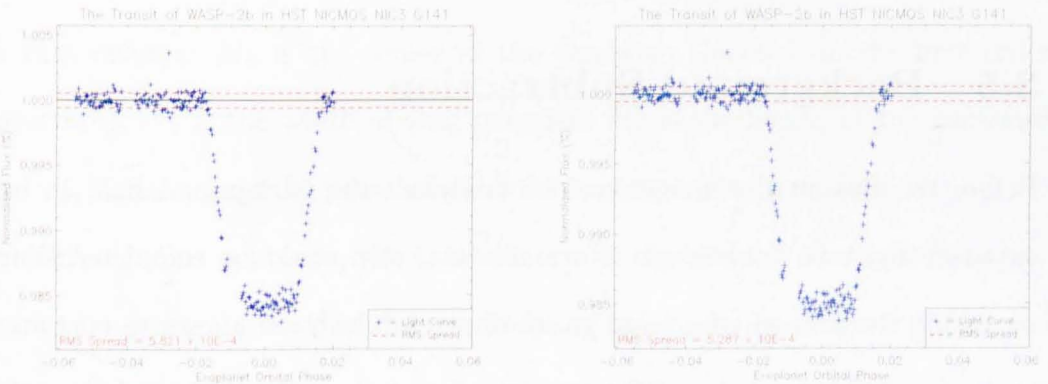
To test the veracity of this choice, both the composite and spliced flat fields were applied to the same data set. In each case, the same processing steps were followed and the same box dimensions were used (see Section 3.6). The resulting light curves showed no discernable differences and negligible variations in RMS Spread (see Section 3.7). Given that there are no discernable disadvantages to using the composite flat field alone - and given the potential inaccuracies caused by the spliced flat field, the decision to use only the composite flat field would seem to be sound.

It should be noted that, by omitting the flat-fielding step, the RMS Spread of the resulting light curve (see Section 3.7) was found to increase across all data sets. However, after applying a process of detrending (see Chapter 4) to both sets of curves, the resulting RMS Spread (see Section 3.7) was either comparable to the flatfielded data or, in some cases, superior - as is demonstrated in Figure 3.7. However, in such cases it was clear from the resulting light curve that, despite the superior RMS, residual systematics remained.

The most likely explanation for this enhancement is that some systematics were enhanced by the omission of the flat fielding process - allowing the de-



(a) Before Detrending



(b) After Detrending

Figure 3.7: *The same transit light curve (blue) from flat fielded images (left) and with flat fielding omitted (right). The upper plots show the transit before the removal of systematic noise; where the RMS spread (red dashed lines) of the field-field-omitted images greater than the flat fielded light curve. In contrast, the lower plots show each light curve after the removal of systematic noise (Chapter 4); where the flat-field-omitted curve exhibits the superior RMS spread. In all plots, black points represent exposures removed from processing (see Section 3.3.3).*

trending process to better model and remove these effects. However, with the flat fielding process omitted, the remaining detrending parameters could not be properly calculated from the data, preventing the removal process (detailed in Chapter 4) from properly accounting for their effect.

However, despite this finding, the flat field is still applied in all cases. This decision was made to provide the most reliable images to the Gaussian Fitting procedure (see Section 3.5), the results of which are later used to define the region encompassing the first order spectrum (see Section 3.6) and in the removal of systematic noise (see Chapter 4).

3.5 Background Subtraction

Before the flux in the images can be analysed, the background light in each exposure must be subtracted. Normally, this step could be completed simply by taking the average of a large group of pixels which are known to contain no background objects and no light from the target star. However, such a method takes no account of the variations in the recorded background across the chip and, while this effect would normally be negligible, in our case the light in each image will be measured from $\sim$ 2500 pixels. Such a large area would be much more likely to contain regions where such factors may occur and, if far enough removed from the First Order Spectrum (FOS) itself, may produce an entirely different background level than that present at the FOS.

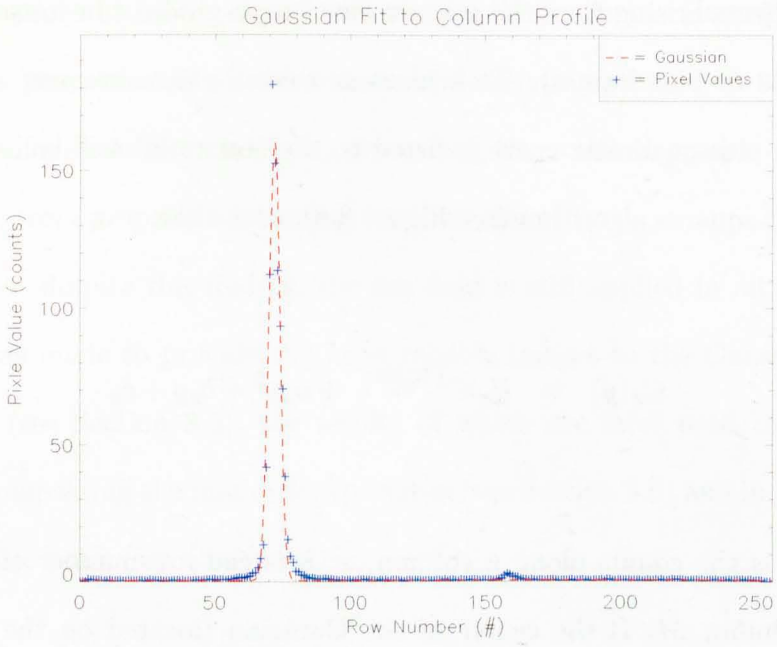
Therefore, in order to be as precise as possible, the background counts in each column are measured individually (thereby accounting for any trends across the

detector, effects arising from the grism's wavelength-gradient or innaccuracies in the flat field or bias frames). To achieve as robust a measurement as possible, a Gaussian-plus-quadratic curve is fitted to the counts in each column using a simple least-squares algorithm (see Figure 3.8), such that:

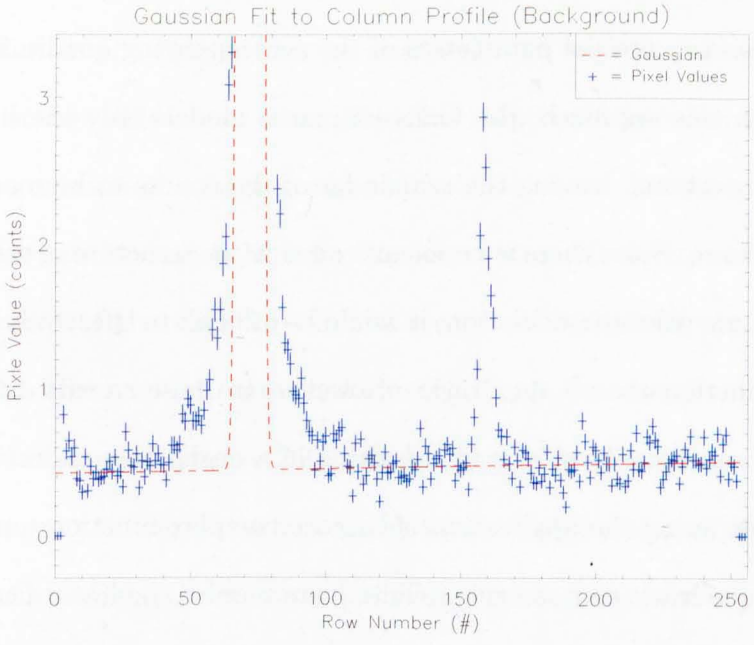
$$C_x(y) = H_x e^{-\left(\frac{y-M_x}{W_x}\right)^2} + a_x y^2 + b_x y + c_x \tag{3.9}$$

where, C_x is the counts along a column, x; y is the row-number of each pixel in that column; M_x is the center of the Gaussian (located on the first order spectrum); W_x is the width of that gaussian; H_x is the height of the gaussian; and a_x, b_x and c_x are the parameters of the accompanying quadratic fit.

Through this approach, the Gaussian curve models only the light from the first-order spectrum, leaving the remainder of the counts to be modelled by the quadratic curve. Since there is no means to know the structure of the background in each column prior to calculation, it would be difficult to effectively apply a more complex function than a quadratic. However, in most cases a linear fit alone provides an acceptably precise fit, resulting in a negligible quadratic component (a_x) and rendering the application of a more complex function unnecessary. In effect, the quadratic component provides insurance on the linear fit - resulting in a more robust estimate of the background light falling directly on the first order spectrum. Effectively, the Gaussian estimates the light from the star while the quadratic models the background.



(a) First Order Spectrum



(b) Background

Figure 3.8: *The pixel values of a typical column in a grism image (blue), selected to bisect the first order spectrum. The red, dashed line shows the best-fit Gaussian-plus-quadratic curve for this column. Plot (a) shows the overall fit to the column, dominated by the Gaussian itself while plot (b) shows the plot to the background in greater detail.*

In order to guarantee a reliable fit to the first order spectrum, an artificial gaussian (created to be 10% of the maximum pixel value in the FOS) is added to every column at a significant distance from the FOS. This ensures that, for columns outside the first order spectrum, the fitting procedure will locate on this artificial peak while still approximating the quadratic fit to the background. For columns covering the first and second order spectra, the artificial curve is insignificant and so the fitting procedure locates on the light from the target object. This approach also allows easy identification of those Gaussians which instead locate on the first order spectrum, which will be useful in later processing steps (see Section 4.2.1.1).

The quadratic curve can then be used to calculate and subtract the background light from each pixel in the column, such that:

$$C'_x(y) = C_x(y) - a_x y^2 - b_x y - c_x \tag{3.10}$$

where, C'_x is the background-subtracted light along a column, x; C_x is the recorded counts in that column; y is the row address of each pixel in that column; and a_x, b_x and c_x are the coefficients of the quadratic fit to the background light in that column.

Repeating this for every column in each image provides a robust means of reliably subtracting the background counts. The parameters of the Gaussian fit are then saved to be later used in the identification and removal of systematic

noise (see Chapter 4).

3.6 Locating the First Order Spectrum

In order to collect the counts in the first order spectrum of each image, its boundaries must first be precisely located. To this end, we define a light-gathering box around the first order spectrum, to ensure the robust treatment of all exposures.

To calculate the best dimensions of this box, the grism images are combined to create a high-S/N composite grism image (as was done with the direct images when locating the target star). This composite image is then collapsed in both the x and y dimensions to create arrays of column- and row-totals. These arrays can then be analysed independently to determine the optimum x and y dimensions for the light-gathering box.

3.6.1 Light-Gathering Box: Spatial Dimensions

To determine the location of the box in the y-axis (also called the 'spatial' axis, since it is perpendicular to the grism's dispersion) a simple Gaussian is fitted to the composite row-totals, such that:

$$C_{Tot}(y) = H_{Tot} e^{-\left(\frac{y-M_{Tot}}{W_{Tot}}\right)^2} \tag{3.11}$$

where $C_{Tot}(x)$ is the total counts in each row, y, in the composite grism image; H_{Tot} is the width of the best-fit Gaussian to the first order spectrum in these

row-totals; M_{Tot} is the center of this Gaussian; and W_{Tot} is the width of the Gaussian.

From this Gaussian, the limits of the box are defined as three standard-deviations (as measured from the best-fit width of the Gaussian) on either side of the Gaussian center:

$$B_{y,Min} = M_{Tot} - 6\sqrt{2\ln(2)}\,W_{Tot} \tag{3.12}$$

$$B_{y,Max} = M_{Tot} + 6\sqrt{2\ln(2)}\,W_{Tot} \tag{3.13}$$

where, $B_{y,Min}$ is the lower boundary of the light-gathering box in the spatial, y, axis; $B_{y,Max}$ is the upper boundary of the box in this axis; M_{Tot} is the center of our best-fit Gaussian; and W_{Tot} is the width of that Gaussian.

A typical example of the spatial boundaries resulting from this process can be found in Figure 3.9.

3.6.2 Light-Gathering Box: Spectral Dimensions

Defining the spectral (x) dimensions of the light gathering box is a more difficult task since the wavelength dispersion of the first order spectrum follows no simple mathematical rule (like the Gaussian used to characterize the box y-dimensions). Instead, the x-dimension of the first order spectrum is described by the wavelength sensitivity function (WSF) of the detector; which gives the relative sensi-

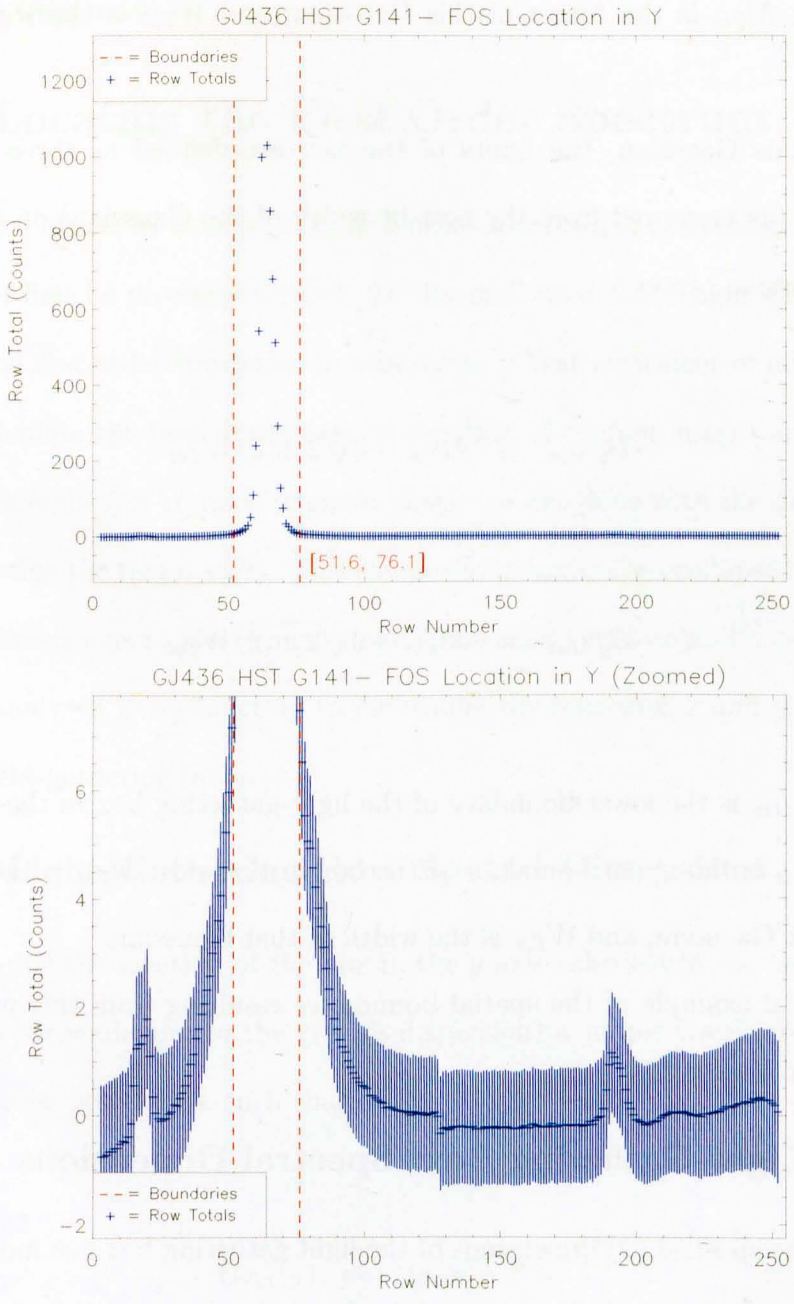


Figure 3.9: *The bolundaries (red) of the light-gathering box in the spatial (y) dimention in relation to the average row-totals of the composite grism image (blue) from which these bounaries were calculated. The upper plot is scaled to the column totals, while the lower plot is scalled to the background of the composite image.*

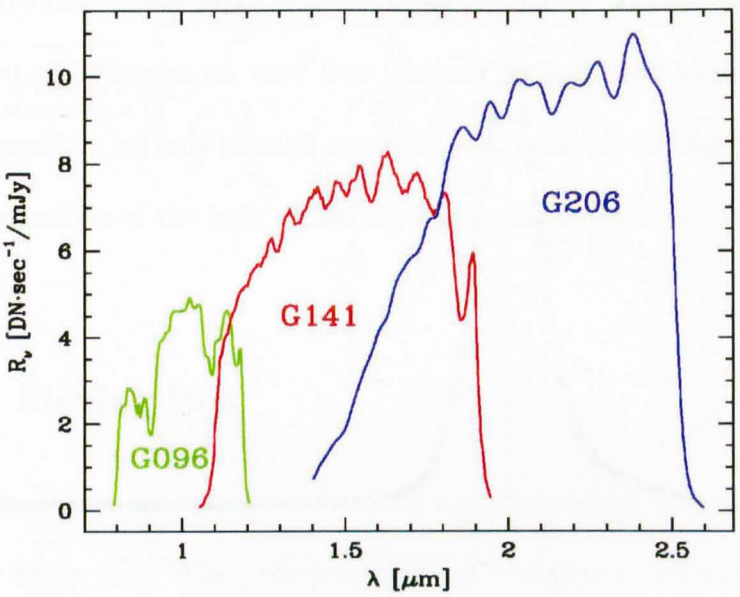


Figure 3.10: *The wavelength sensitivity functions of the three NICMOS grisms - G096 (green), G141 (red) and G206 (blue).*

tivity of the detector to various wavelengths of light with the same incident flux. Figure 3.10 shows the wavelength sensitivity functions for the ranges covered by the three NICMOS grisms.

Since there is no set of parameters governing the WSF (as there would be for a Gaussian or similar function) we must define our own model by which to find the best-fit of the WSF to the column totals of the composite grism image. If we assume zero background light (a safe assumption, since the background has already been subtracted in Section 3.5) we can find the best approximation of the WSF by normalizing it to have the same area under the curve as the FOS.

This normalized WSF can then be fitted to the first order spectrum by finding the position in the x axis with the minimum χ^2 goodness-of-fit statistic relative

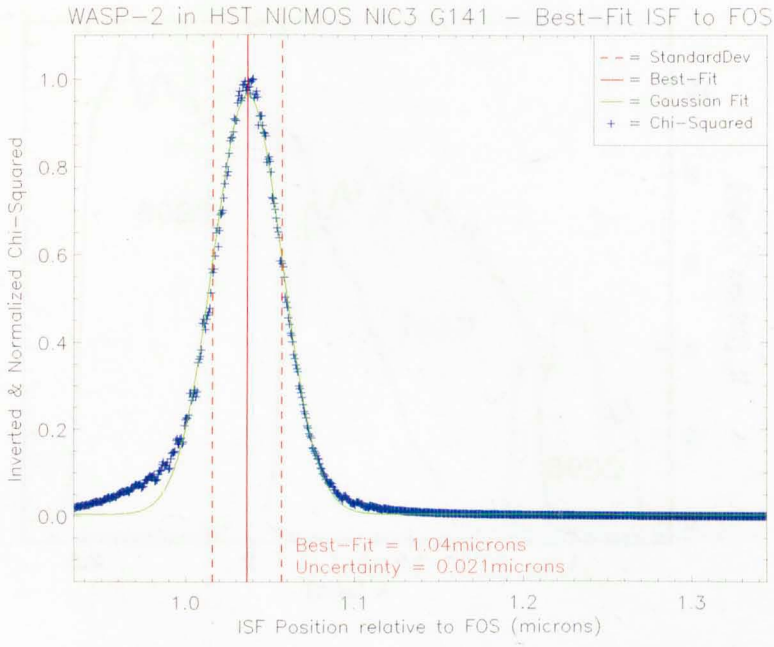


Figure 3.11: *The χ^2 goodness-of-fit statistic for a model dispersion relation compared to an array of column-totals for a composite grism image, at a series of wavelength offsets ($\delta\lambda$). The χ^2 values shown (blue) have been inverted and normalized to show the best-fit at the peak of the plot. The overplotted Gaussian curve (green) shows how the center of this chi-squared distribution has been measured and used to locate the position of the first order spectrum in the image. This estimated position is denoted by the solid red line, with the dashed red lines showing the estimated uncertainty from that same Guassian.*

to the column totals of the composite image. To find this best-fit position, the χ^2 test is performed with the normalized WSF at various intervals across the first order spectrum. Fitting a Gaussian to the resulting χ^2 distribution (see Figure 3.11) gives a robust measurement of the minimum χ^2 value and the best fit position of the WSF.

Once its position in the spectral (x) axis has been determined, the dimensions of the first order spectrum can be calculated from the length of the wavelength

sensitivity function. This gives a more robust estimate than the length of the FOS itself which is known to vary (see Chapter 4) and, due to its steep sides, would be dependent on only a small number (~ 4) columns on the detector. The resulting dimensions of the light gathering box in the spectral, x, axis are shown in Figure 3.12.

3.6.3 A Moving Box

For each light curve, the same box location and dimensions were used for every image in the same visit. When the possibility of applying a different box to each image was investigated, this was found to greatly detriment the resulting light curve (see Figure 3.13).

There are likely several causes for such a detriment. Firstly, the motive behind the application of a moving box is based on the knowledge that the first order spectrum moves slightly during each visit - introducing a great deal of systematic noise. However, it is likely that a great deal of this noise is caused by variations in pixel sensitivity across the first order spectrum, rather than the outer edge of the FOS moving in or out of the box. Since the later step to remove systematic noise (see Chapter 4) is expected to account for both of these factors, there is no detriment for using a static box over a moving one.

Secondly, the dimensions of a moving box must be re-calculated from each image. This removes the advantage of using the high-S/N composite grism image and makes the resulting dimensions of each box less precise and far less robust.

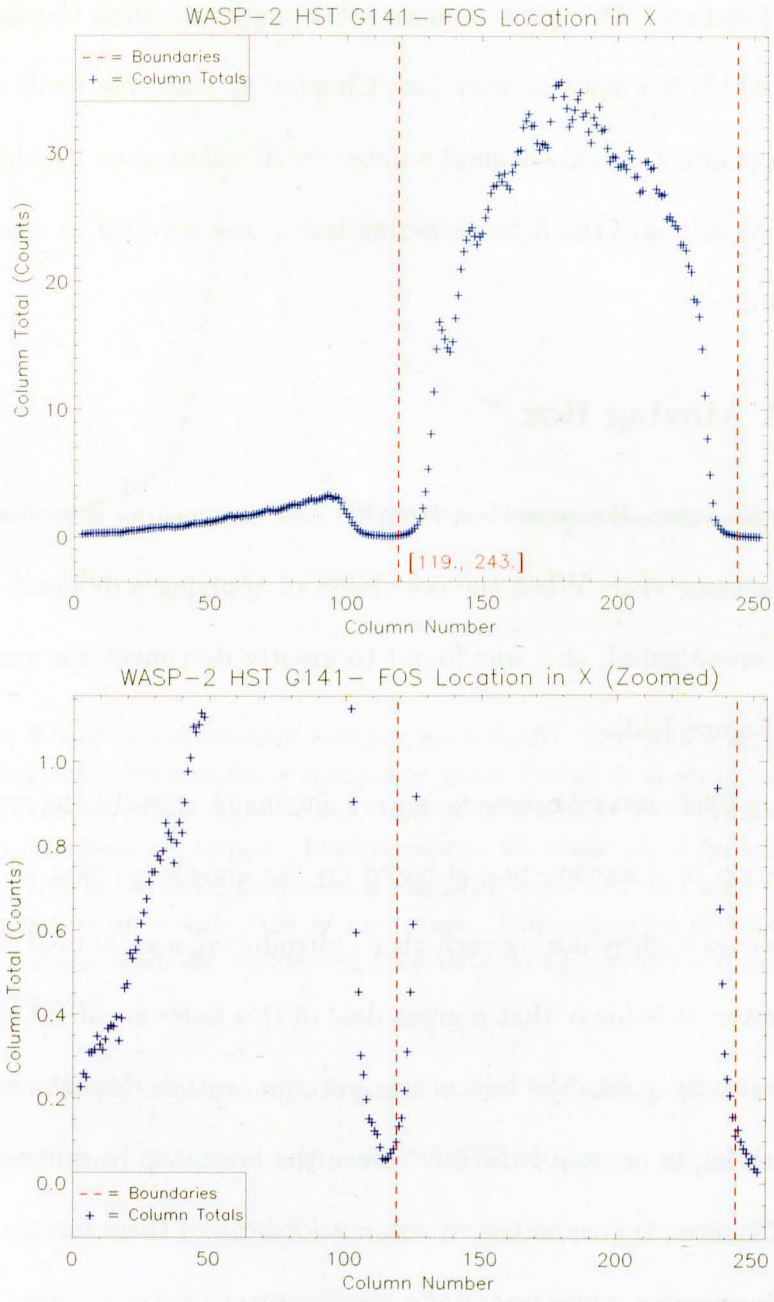


Figure 3.12: *The boundaries of the light gathering box (red) in the spectral (x) dimension shown in relation to the averaged column totals of the composite grism image from which the boundaries were calculated (blue). The upper plot is scaled to the column totals, while the lower plot is scalled to the background of the composite image.*

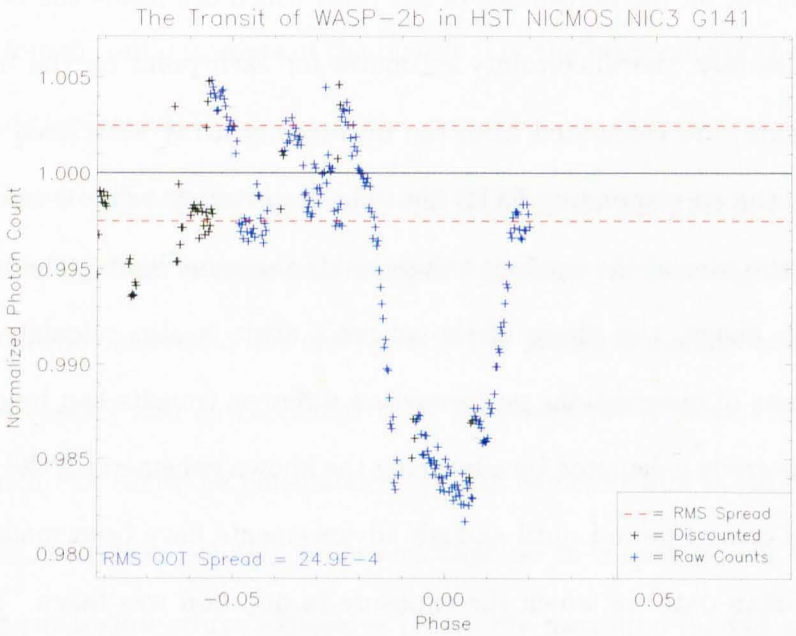


Figure 3.13: *A transit light curve (blue) created through application of a moving box whose dimensions were calculated by means described in Chapter 4. The impact of systematic noise on the data is readily apparent in both the features of the light curve and the out-of-transit RMS spread; denoted by the dashed red lines. The black points are those exposures previously flagged to be omitted from future calculations.*

3.7 Creating the Raw Light Curve

Once a robust light-gathering-box has been defined, the counts in each image can be totalled simply by adding together the value of every pixel in that box. Any pixels which fall partly inside the box are included as a percentage value of their total, dependent on the percentage of the pixel which lies inside the box.

Simultaneously, the uncertainty estimates for each point on the light curve (each exposure) are calculated from the uncertainty array associated with each exposure in the corresponding FITS file. The uncertainty value is calculated as the root of the sum of the squared values inside the same light-gathering box.

For each image, the phase of the planet's orbit is also calculated so that different visits of observations performed on different transits can be combined. The phase is easily calculated by advancing the known ephemeris of the transit by the planet's orbital period until enough advancements have been made to reach the same Julian date[1] in which the exposure in question was taken. The phase is normalized so that the center of the transit stands at 0 and the extremes at ± 0.5. As such:

$$\phi(l) = \frac{t(l) - (nP + E)}{P} \tag{3.14}$$

where $\phi(l)$ is the exoplanet orbital phase at the time of an exposure, i; $t(l)$ is the time at which that exposure was taken, expressed as a Julian Day; E is the julian

[1]Both the Julian date of the exposures and the Julian date of the exoplanet's Epoch, E, were converted into Modified Heliocentric Julian Days prior to any calculations.

date of the exoplanet's epoch; n is the number of orbits completed by the planet since that ephemeris was recorded; and P is the length of that period in days.

3.7.1 Locating the Transit

Despite the systematic noise, it is possible to arrive at a rough estimate for the first and fourth contact-points of the transit (i.e. the beginning of the ingress and the end of the egress).

The center of the ingress and egress can be located simply by noting the points at which the light curve crosses its median flux value. Depending on the number of points in the data set, a line can be fit to the points immediately around these estimated centers and extrapolated until it crosses the a mean Out-Of-Transit value which can easily be calculated from the recorded flux in orbits known to include no in-transit points. To ensure that no in-transit points fall into our out-of-transit regime, three exposures before the measured ingress and after the egress are considered in-transit in subsequent calculations.

Once our out-of-transit points are known, the light-curve can be normalised to set the mean OOT flux to 1 so that the change in flux during transit is presented as a fractional change in this normalized flux value. By calculating the RMS spread of the OOT points about this mean value, it is possible to measure the impact of noise on the light curve. This RMS spread value will be useful when analysing the effectiveness of the systematic noise reduction covered in Chapter 4.

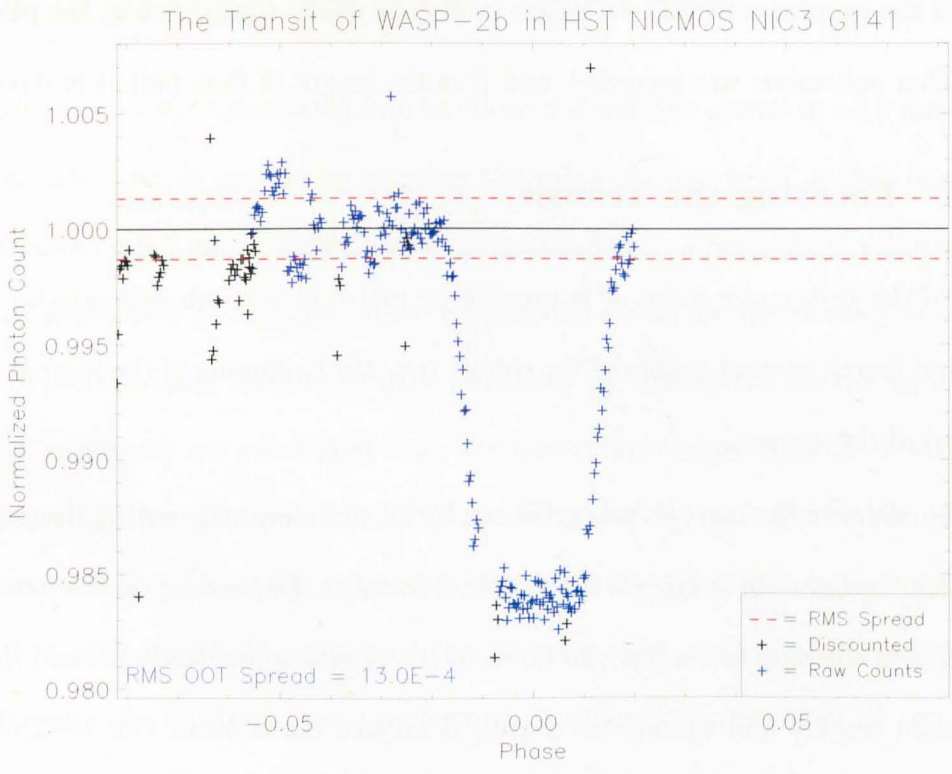


Figure 3.14: *A typical raw transit light-curve. Note the prevalence of systematic effects, particularly in the early Out-Of-Transit (OOT) orbits.*

3.7.2 The Light Curve

Plotting all visits against the calculated phase gives a raw transit light curve, like that seen in Figure 3.14. At this stage, the NICMOS light curves are typically plagued by systematic effects which must be removed if precise planetary parameters are to be derived.

Chapter 4

Methods 2 - Systematic Noise

4.1 Chapter Outline

The following chapter details the identification and removal of systematic noise applied to the raw transit light curve. A review of the sources of systematic noise - together with how each source is parameterised - is followed by a comparative analysis of the means by which such noise can be robustly removed: a traditional (in the field) method of linear regression versus a new, iterative technique termed 'Preferential Decorrelation'. The chapter concludes with various means by which any residual systematic noise in the processed light curve can be measured - specifically by the binning, 'prayer-bead' and wavelet techniques.

4.2 Sources of Systematic Noise

The NICMOS instrument has various sources of systematic noise. Usually, such noise would have a small or negligible effect on results. However, since our observations require such high precision, these systematic effects can have a significant impact on results.

Perhaps the most significant source of systematic noise are pointing variations which occur over each orbit of any HST visit of observations. As the telescope moves slightly during the course of each orbit, the location of the target object on the detector moves as well. As a result, the light falling on the detector progressively moves from one pixel to the next - imprinting any variation in pixel sensitivity onto the resulting light curve.

If this systematic noise is to be removed, it must first be parameterised in order to discern its effect on the measured flux. Thankfully, we do not need to parameterise the pixel sensitivity itself, nor any other direct systematic noise source. Instead, we can simply model parameters which are themselves representative of the systematic effects, yet independent of flux.

4.2.1 Pointing Variations

Most significantly, the movement of the first order spectrum across the detector must be characterised using various parameters. Primarily, we must model the movement in both the spectral and spatial axes. However, in addition to this, the first order spectrum is also known to vary its width (in the spatial axis) with both wavelength and time, and, independently, its length (in the spectral axis) with time. Furthermore, the orientation of the first order spectrum will change throughout each visit - effectively causing the first order spectrum to rotate over time. All of these parameters must be measured if the systematic effects caused by such pointing variations are to be characterised and removed from the light curve.

Such measurements are performed with the Gaussian curves fitted to each column while measuring the detector background (see Section 3.5). Since any variations in the position of the first order spectrum will be reflected by changes in these Gaussian fits, we can use these curves to parameterise any pointing variations, even if the Gaussian itself does not perfectly model the flux.

Each column in the first order spectrum has its own Gaussian, with the equation:

$$G_x(y) = H_x e^{-\left(\frac{y-M_x}{W_x}\right)^2} \tag{4.1}$$

where G is the Gaussian; x is the column-number in the current image; y is the row number in the current column; H is the height of the Gaussian; M is the center of the Gaussian; and W is the width of the Gaussian.

We can therefore model the position of the first order spectrum with the following parameters:

4.2.1.1 Spatial Position (Y-Position)

The location (P_0) of the first order-spectrum in the spatial axis can be simply derived from the Gaussian-centers of the fits to each column. For each image, the Spatial Position is defined as the average Gaussian-center of every column crossing the first order spectrum.

In order to ensure a robust measurement, only columns where the Gaussian fit has located squarely on the first order spectrum are included in this average. The boundaries of this region are determined thusly:

1. Define a region of thirty columns in the center of the first order spectrum (typically the middle thirty columns of the light gathering box).

2. Calculate the Standard Deviation of the Gaussian-centers in that region.

3. Expand this region by one column on each side along the first order spectrum. If the Gaussian-center of either new column falls more than three standard deviations from the mean of the current region, that column is rejected. Otherwise, that column is included in the region and the standard deviation is recalculated.

4. Repeat step 3 until three adjacent columns are rejected in either direction.

5. Set the last accepted column as the edge of the first order spectrum in this image.

6. Repeat step 3 until three adjacent columns are rejected in the remaining direction.

7. As before, set the last accepted column as the other edge of the first order spectrum in this image.

8. Move the boundaries of this region ten pixels towards the center of the first order spectrum - thereby ensuring that the columns at the edge of the first order spectrum (where a smaller number of counts would be expected to result in a less reliable Gaussian fit) are not included in subsequent calculations.

With a robust region defined, the spatial position of the first order spectrum in this image is calculated as:

$$P_0 = \frac{1}{k-j}\sum_{l=j}^{k} C_l \tag{4.2}$$

where P_0 is the spatial position; l is the column number; j is the lower limit of the region containing the first order spectrum; k is the upper limit of this region; and C_l is the Gaussian-center for the column l.

The user can change the number of standard deviations used to calculate the acceptance threshold using the Config file. The same applies for the number of columns excluded from the edges of the first order spectrum.

The typical Gaussian-center values for all columns in an image can be seen in Figure 4.1. Additionally, a typical array of spatial positions across a complete visit of observations, can be seen in Figure 4.2.

4.2.1.2 Spectral Position and Length

The position of the first order spectrum in the spectral axis (P_1) is calculated via the same method as the boundaries of the light gathering box (as detailed in Section 3.6).

However, while the light gathering box is calculated from a high S/N composite image, the spectral position must be calculated independently for each image. Unfortunately, the far lower S/N of each image's individual column totals provides a poor reference for the Inverse Sensitivity Function (ISF) to centre on. As such, it is preferable to use the heights of the Gaussian curves fitted to each

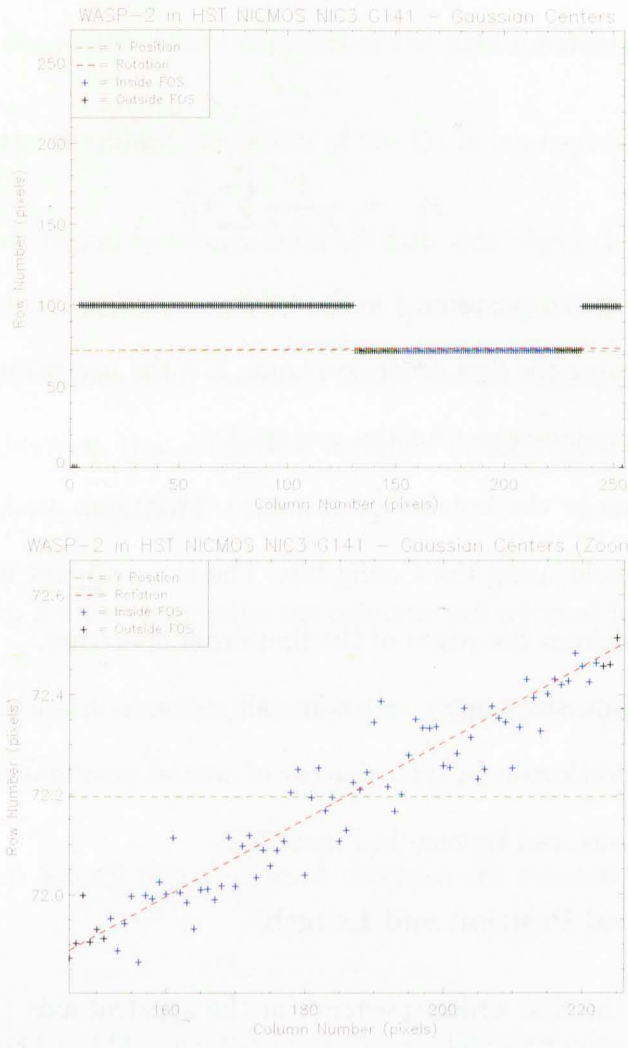


Figure 4.1: *The pixel-values on which the best-fit Gaussian curves are centered for the columns in a single image. The values in blue are those values included in the calculation of the subsequent Spatial (y) Position (P_0, dashed red line) of the first order spectrum in this image and the Rotation of the first order spectrum (P_5, dashed green line). The upper plot shows the full range of Gaussians across the detetor. The lower plot has been zoomed to show only those points in and around the first order spectrum.*

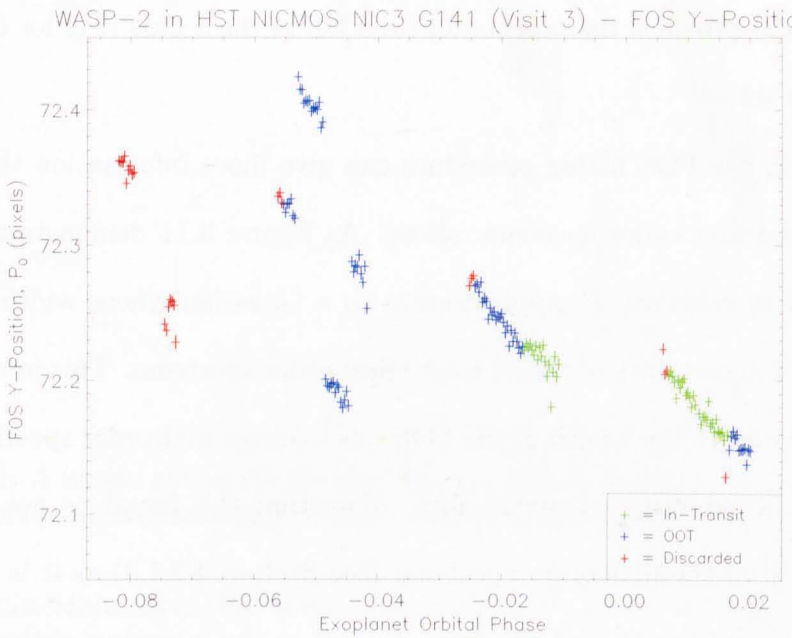


Figure 4.2: *A typical array of first order spectrum Spatial Positions for a single HST visit of observation. The points shown in red are from images flagged to be excluded from calculations - typically due to a large quantity of anomalous pixels or an expected large amount of systematic noise. The points shown in blue are calculated from exposures taken outside of the expolanetary transit, while points shown in green are from exposures taken while the planet was expected to be in transit. The gaps in the curve are the result of HST orbit breaks where the telescope passed behind the Earth, relative to the target object.*

column to represent the profile of the FOS in the x-axis. The resulting profile is far more coherent than the column totals and has already undergone screening for anomalous values as part of the earlier Gaussian fitting process. Performing the FOS fitting procedure on the Gaussian heights of each image (as detailed in Section 3.6.2) provides the position of the FOS in the x-axis (P_1) for each image in the data set.

However, the FOS fitting procedure can give more information than the location of the first order spectrum alone. As Figure 3.11 demonstrates, the χ^2 curve used to calculate P_1 approximates to a Gaussian whose width relates to the range of uncertainty of the fit to the first order spectrum. This uncertainty is directly related to the length of the FOS - as a longer first order spectrum would result in a wider range of uncertainty. Note that this length is not analagous to the width of the first order spectrum (see Section 4.2.1.3) as it is dependent on the state and location of the grism as well as the focussing of the telescope. Furthermore, since the defocussed point spread function of the target is known to be asymmetric, it is unlikely that variations in the width of the PSF would be the same in the spatial and spectral axes. Recording the width of each χ^2 Gaussian gives another parameter by which to characterise the position of the FOS - namely the "Spectral Length", P_4.

A typical plot of the Spectral Position (P_1) of the first order spectrum over the course of a single visit of observation is shown in Figure 4.3. Correspondingly, Figure 4.4 shows the Spectral Length (P_4) of the first order spectrum over the same visit of observation. The uncertainties on each parameter are simply calculated

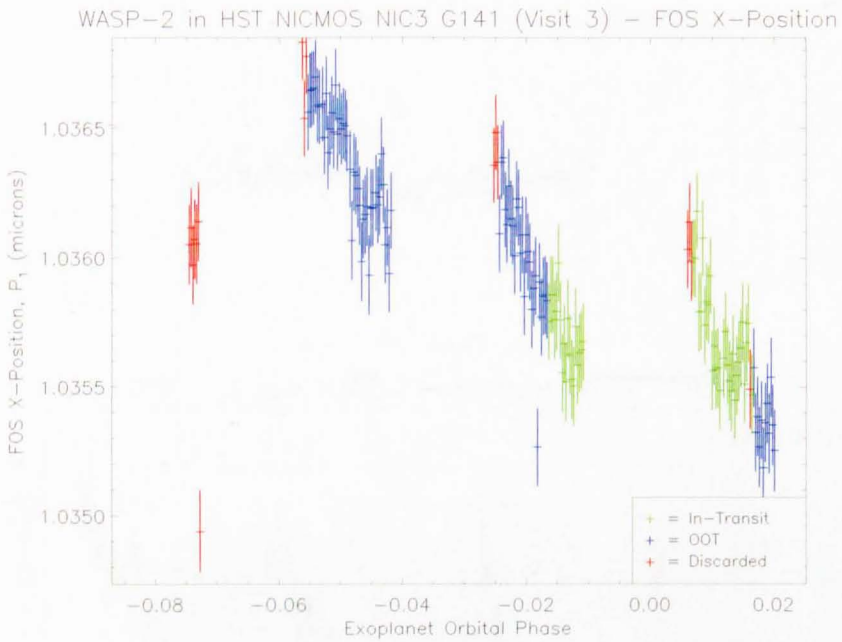


Figure 4.3: *A typical plot of the Spectral Position of a grism first order spectrum over a single HST visit of observation. Points on the graph are coloured as in Figure 4.2.*

from the uncertainties of the Gaussian coefficients from which the parameters are derived.

4.2.1.3 Spatial Width and Tapering

Variations in the width of the first order spectrum (P_2) measured in the spatial (or x) axis arise from focussing variations over the course of a visit of observation. The width is calculated from the corresponding widths of the Gaussian fits to each column (W_x), as shown in Figure 4.5. However, the width of the first order spectrum typically varies along its length, either widening or narrowing with increasing wavelength. Therefore, a simple average is not guaranteed to be representative of the variation in width from image to image.

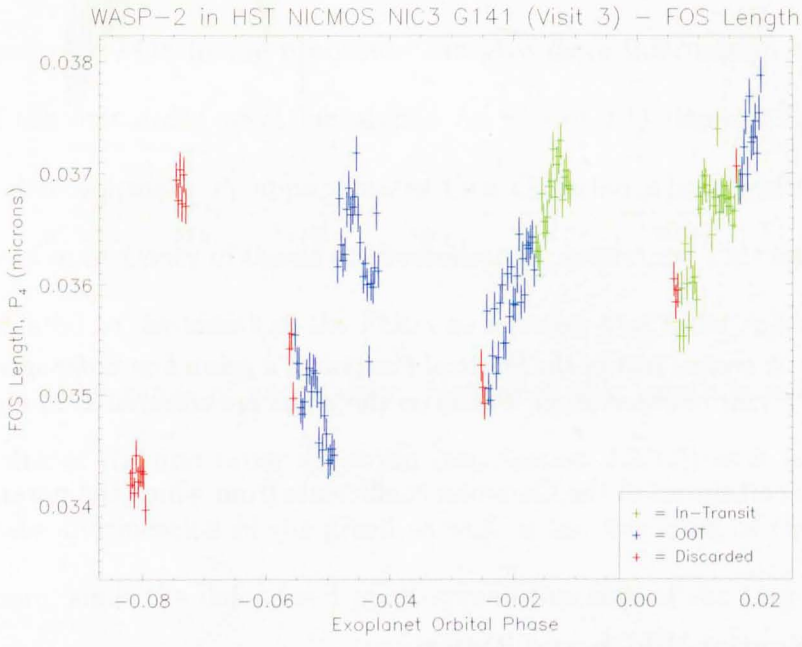


Figure 4.4: *A typical plot of the Length of a grism first order spectrum over a single HST visit of observation. Points on the graph are coloured as in Figure 4.2.*

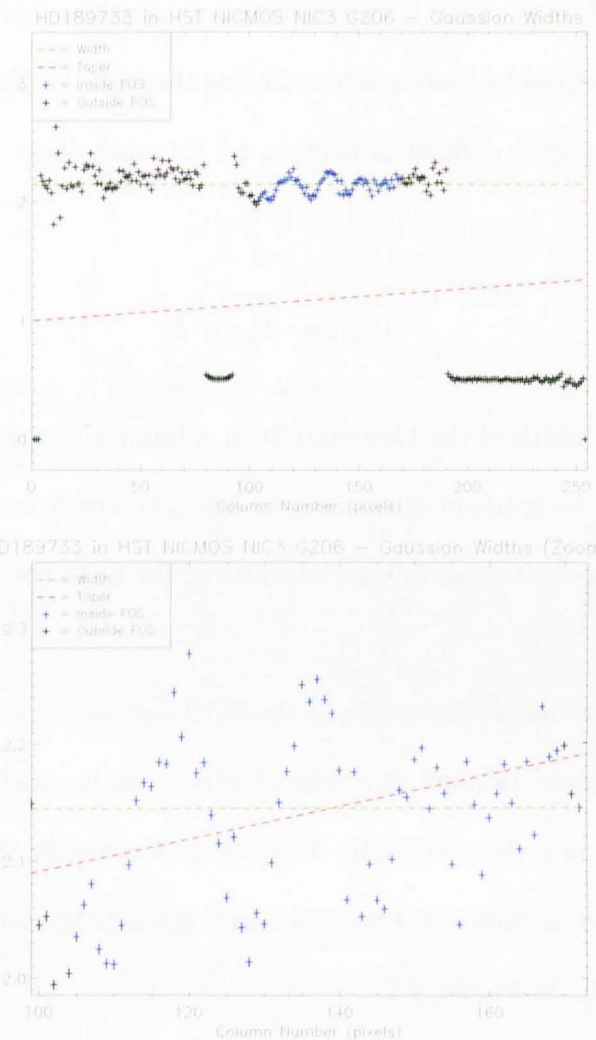


Figure 4.5: *The widths of a series of Gaussian curves (W_x) fitted to the columns of a single HST grism image. The points plotted in blue are those defined (in Section 4.2.1.1 as being firmly inside the range of the first order spectrum. The overplotted, green, dashed line demonstrates the measurement of the width of the first order spectrum (measured as the intersect of this line with the y-axis) The overplotted red line (offset from position of FOS for clarity) represents the tapering of the first order spectrum (measured as the gradient of this line). The upper plot shows the full range of Gaussians across the detetor. The lower plot has been zoomed to show only those points in and around the first order spectrum.*

To avoid this degeneracy between temporal variation and spectral variation, the width is measured by finding a best-fit line through the Gaussian width values across the same region defined in Section 4.2.1.1, such that:

$$W_x = P_3 x + P_2 \tag{4.3}$$

where W_x is the width of the Gaussian fit in column x (corresponding to a wavelength, λ); P_3 is the gradient of the best-fit line (effectively measuring the taper of the first order spectrum); and P_2 is the width of the fit to the first order spectrum at $x = 0$.

In this way, we model the width of the FOS and any tapering along its length as separate variables (P_2 and P_3) both of which can be used to characterise the location of the first order spectrum. A typical plot of the Spatial Width evolution over a single visit is shown in Figure 4.6 and the corresponding plot of Spectral Tapering is shown in Figure 4.7.

4.2.1.4 Rotation

All previous parameters have made a precise distinction between the spatial (y) and spectral (x) axes in each image. However, this makes the erroneous assumption that the applied grism will be properly aligned for each exposure. In reality, the grism is seldom perfectly aligned to be completely parallel with the ideal spectral axis and, furthermore, its alignment is prone to variations during each visit of observations.

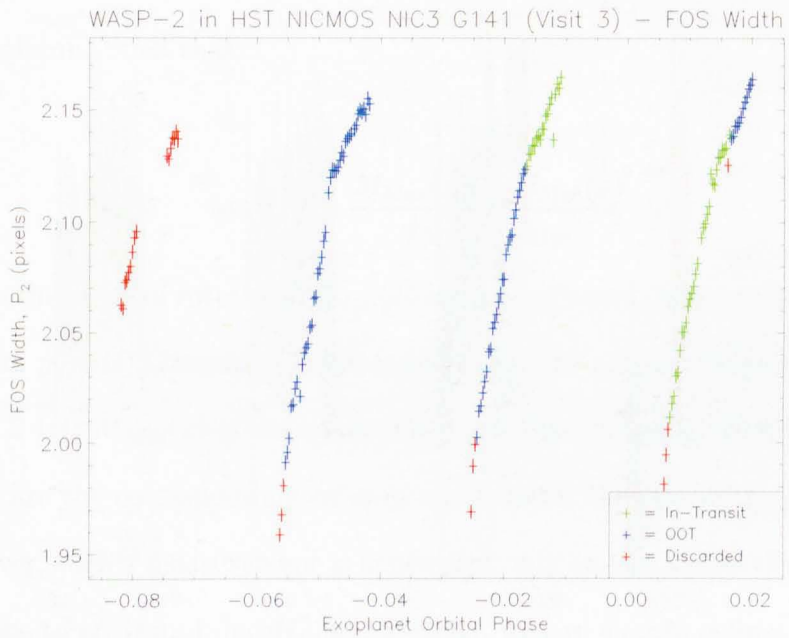


Figure 4.6: *A typical plot of the Width of a grism first order spectrum over a single HST visit of observation. Points on the graph are coloured as in Figure 4.2.*

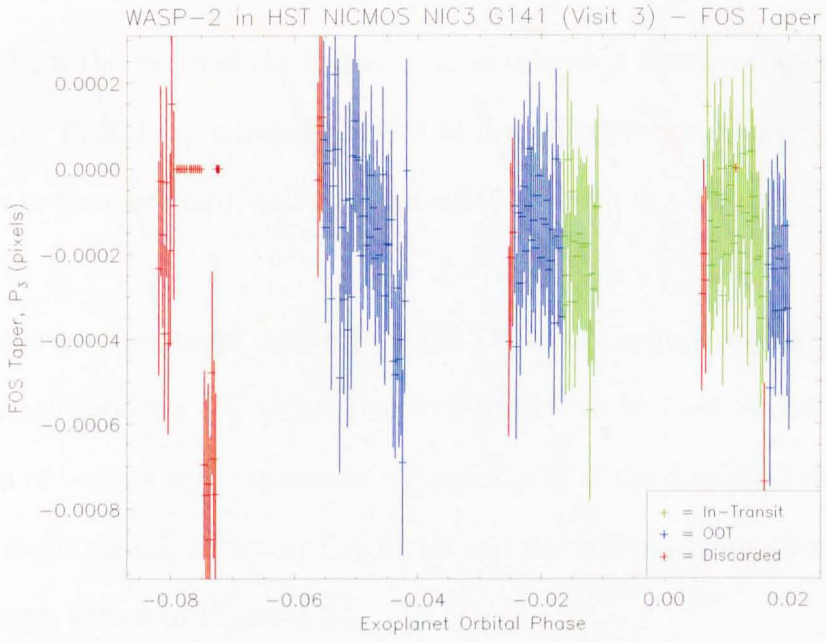


Figure 4.7: *A typical plot of the Tapering of a grism first order spectrum over a single HST visit of observation. Points on the graph are coloured as in Figure 4.2.*

In order to parameterise these variations and properly account for their impact on the resulting light curve, we must robustly measure the Rotation of the first order spectrum in each image - defined as the deflection of the effective center of the first order spectrum from its ideal position parallel to x and perpendicular to y.

This parameter (P_5) can be calculated from the centers of the Gaussian Fits to each column. Such that:

$$\tan\theta = \frac{M_{Max}(x) - M_{Min}(x)}{x_{Max} - x_{Min}} \tag{4.4}$$

where θ is the angle of rotation; $M_{Max}(x)$ is the maximum (highest row-number in the spatial, y, axis) Gaussian Center taken inside the column-range calculated in Section 4.2.1.1; $M_{Min}(x)$ is the equivalent minimum Gaussian Center; and x_{Max} and x_{Min} are the corresponding column numbers for the two values of M.

However, such a measurement is dependent only on two values of C and likely to be prone to error and uncertainty. Instead, we can simply define:

$$P_5 = \frac{\Delta M(x)}{\Delta x} \tag{4.5}$$

which has the form of a gradient. As such, we can calculate a far more robust measurement of the rotation by locating a best-fit line to the complete array of Gaussian column centres inside the previously calculated range. Note that P_5 has simply been defined as $\tan\theta$ rather than simply θ since this is the same quantity given more directly.

As such, we can calculate a best-fit line such that:

$$C_x \approx P_5(x - x_{Mid}) + P_0 \tag{4.6}$$

where C_x is any given Gaussian center fitted to a column, x, within the predefined range; P_5 is the effective rotation of the first order spectrum; P_0 is the, already calculated, spatial (y) position of the first order spectrum in the image; and x_{Mid} is the column number of the centre of the first order spectrum, corresponding to the intersect of the FOS with P_0.

Note that the best-fit line could have been defined to have its own constant term independent of P_0. However, on inspection these two terms were found to be almost degenerate - following the same distinct pattern, but with the intercept version calculated in this section having a greater spread. Therefore, this secondary calculation was deemed unnecessary, as the previously calculated parameter better represents the same quantity.

A typical plot of the change in first order spectrum rotation over the course of a visit of observation can be found in Figure 4.8.

4.2.2 Further Parameters

In addition to pointing variations, systematic effects can arise from any factors influencing the grisms, the NIC3 camera or the NICMOS instrument itself. Thankfully, most of such influences are small or vary very little over the timescale of a single visit. However, some such effects are not so obliging and many can not

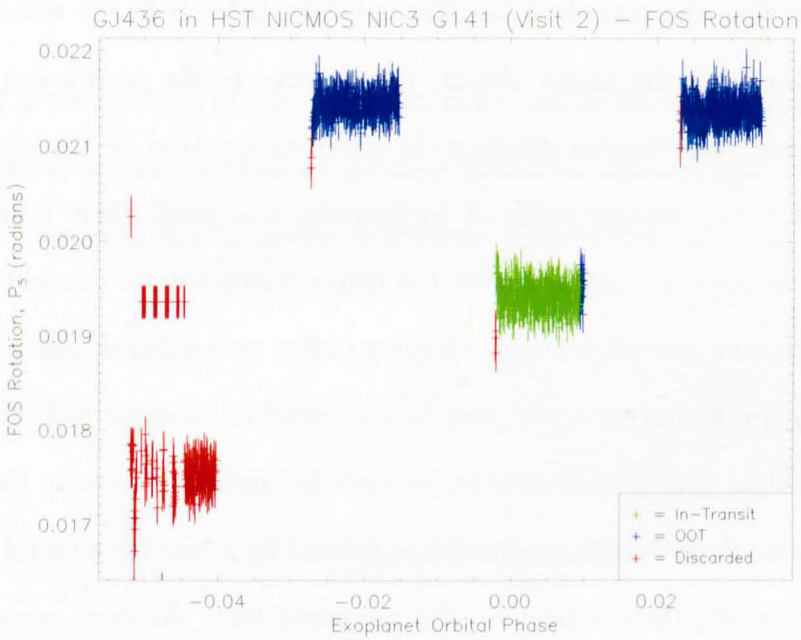


Figure 4.8: *A typical plot of the Rotation of a grism first order spectrum over a single HST visit of observation. Points on the graph are coloured as in Figure 4.2.*

be measured directly. Therefore, we must find further parameters by which to characterise systematic effects.

4.2.2.1 Background

While the background light in each image has already been quantified and removed by this stage of processing (see Section 3.5), this subtraction made the assumption that the measured background light in each image was solely the result of flux from the target object; inaccuracies in the instrument bias and flat-fields; and other factors affecting the electrons reported by any given pixel.

In reality, this measurement of background was most likely influenced by differences in pixel sensitivity across the chip - which can be brought about by variations in temperature, current, bleeding of flux into adjacent pixels, the over-exposure of prior images or, as will be discussed in Chapter 8, detector non-linearity. While only one of these factors can be measured directly (see Section 4.2.2.2), their effects on the measured background light provide a useful paradigm for approximating their impact on the measured flux. As such, we can define another parameter, P_6, as the background light across the first order spectrum in each image.

As previously discussed in Section 3.5, the background light in each column is modelled along with the Gaussian fits to the first order spectrum, such that:

$$B_x(y) = a_x y^2 + b_x y + c_x \tag{4.7}$$

where $B_x(y)$ is the background light in a column, x, at the pixel crossing row y

in the image; and a_x, b_x and c_x are the parameters of the quadratic fit to the background light in that column.

From this, we can calculate the background light at the first order spectrum from the previously calculated spatial (y) position of the FOS, such that:

$$B_x(P_0) = a_x P_0^2 + b_x P_0 + c_x \tag{4.8}$$

where P_0 is the position of the first order spectrum in the spatial (y) axis.

Therefore, we can calculate an estimate of the background light level for each image by averaging $B_x(P_0)$ across all columns in the FOS, such that:

$$P_6 = \frac{\sum_{x=j}^{k-1} B_x(P_0)}{k-j} \tag{4.9}$$

$$P_6 = \frac{\sum_{x=j}^{k-1} (a_x P_0^2 + b_x P_0 + c_x)}{k-j} \tag{4.10}$$

where P_6 is the background flux across the first order spectrum in a given image; j and k are the limits of the first order spectrum in the range as defined in Section 4.2.1.1; a_x, b_x and c_x are the parameters of the best-fit polynomial to the background flux in a column, x; and P_0 is the center of the first order spectrum in the image, as calculated in Section 4.2.1.1.

In most cases, a linear fit to the background is more than sufficient to produce a precise model; however, since the additional processing time required to apply a quadratic model is minimal, there is no real drawback to applying such a model

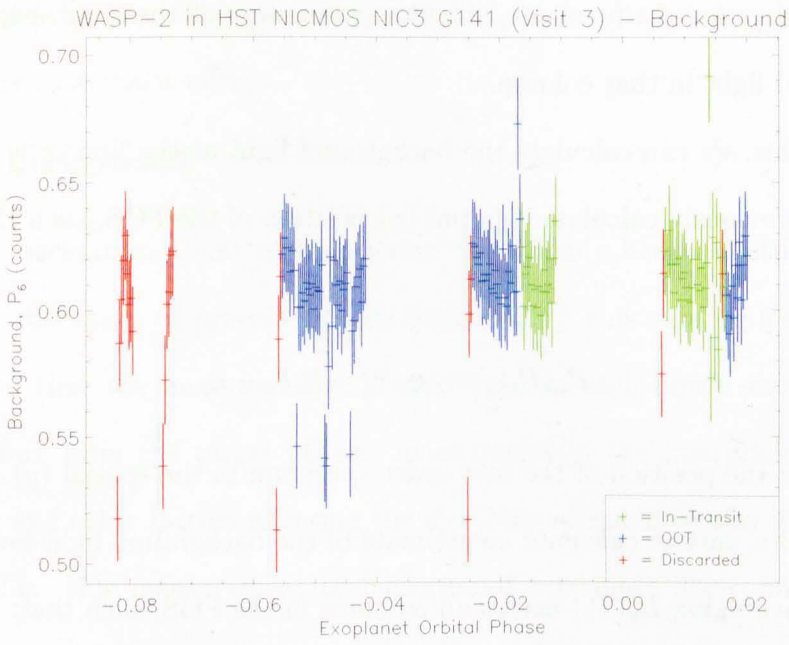


Figure 4.9: *A typical plot of the variation in average Background Flux across the first order spectrum over a single HST visit of observation. Points on the graph are coloured as in Figure 4.2.*

for occassions when a quadratic fit may be required. Typically, the value of a is small enough to be considered negligible. A typical plot of the background flux across the first order spectrum for a single visit of observation can be found in Figure 4.9.

4.2.2.2 Temperature

As noted in Section 4.2.2.1, the temperature of the NICMOS instrument - specifically that of the NIC3 detector - is expected to have a direct influence on the pixel sensitivity of the detector. As such, changes in temperature, P_7, will likely cause corresponding systematic effects.

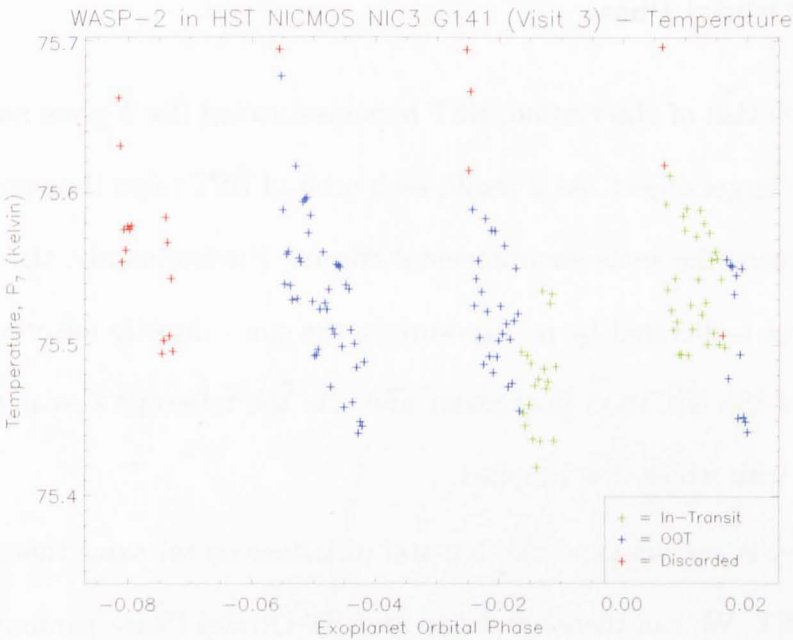


Figure 4.10: *A typical plot of the estimated Temperature of the NICMOS NIC3 detector over a single HST visit of observation. Points on the graph are coloured as in Figure 4.2.*

Thankfully, a good estimate of the detector temperature in each image is provided in the header information of the corresponding `FITS` file. This estimate is, in itself, an aggregated parameter - estimated from the bias voltage across the chip which is known to be directly dependent on the temperature of the chip and the known current supplied to the NICMOS instrument (Viana et al., 2009).

A typical plot of the temperature of the NIC3 detector chip-set for a single visit of observation can be found in Figure 4.10.

4.2.2.3 Orbital Phase

During each visit of observation, HST remains directed (for a given margin of error) at the target object. As a result, each orbit of HST takes the same path and will experience the same environmental effects. Predominantly, the telescope's environment is dictated by its exposure to the sun - directly affecting the temperature of the NICMOS instrument and, via the telescope's solar panels, the electricity with which it is supplied.

As a result, we can expect systematic variations on the same timescale as the orbit of HST. We can therefore define an HST Orbital Phase parameter (P_8) to model such changes.

This parameter is easily calculated as the time expired between the current exposure and the first exposure in the current orbit, such that:

$$P_8 = t_{i,Orb} - t_{0,Orb} \tag{4.11}$$

where P_8 is the Orbital Time; $t_{i,Orb}$ is the exposure date of the current grism image; and $t_{0,Orb}$ is the exposure date of the beginning of the current orbit[1].

A typical plot of the Orbital Time parameter over a single visit of observation can be found in Figure 4.11.

[1]Not to be confused with the exposure date of the first image in the orbit, since guide-star acquisition will take longer in the first orbit. Furthermore, the first image in different orbits is not always guaranteed to be a grism exposure and may be one of the direct images.

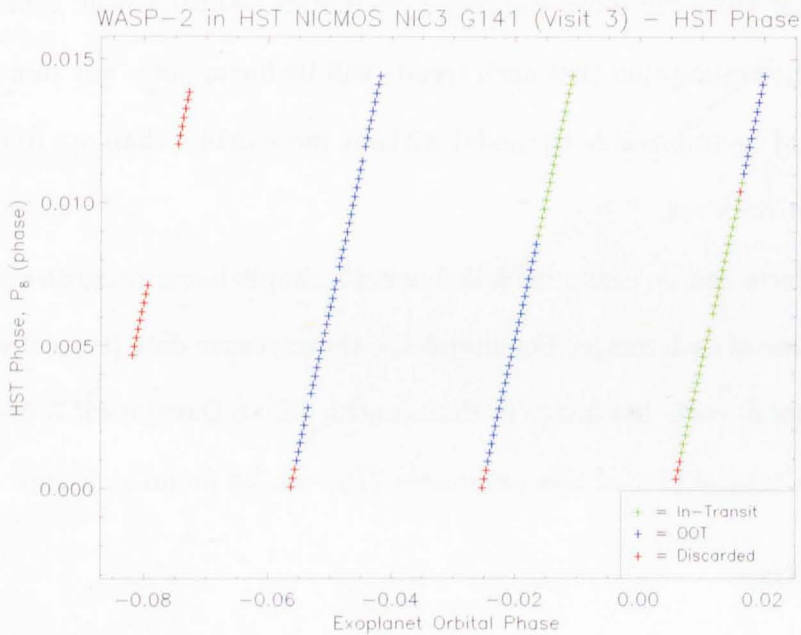


Figure 4.11: *A typical plot of the Orbital Time of each exposure over a single HST visit of observation. Points on the graph are coloured as in Figure 4.2.*

4.2.2.4 Temporal Variations

The only remaining means to parameterise source of systematic noise in the NIC3 detector is to isolate those sources which have a distinct, long-term effect on results. While such factors cannot be parameterised directly, such long-term (i.e. on timescales greater than one orbit) effects would likely result only in long-term trends rather than the more complex systematics resulting from other sources. We make the assumption that such trends will be linear since any more complex trends would be impossible to model without more orbits than are found in any HST transit data set.

Such effects can be easily modelled using a simple linear progression with the exposure time of each image. For simplicity, the exposure date (typically provided in the header of each .fits image in Heliocentric Julian Days) itself is used for this purpose. A typical plot of this parameter (P_9) can be found in Figure 4.12

4.2.3 Bias

As noted by Burke et al. (2010), the bias voltage applied to each chip in the NIC3 camera is known to vary between exposures; resulting in slightly different reading when the bias frame is subtracted by the HST data processing pipeline. The differences between the four chips comprising the NIC3 detector result in each exposure falling into one of seven 'states' determined by the balance of the bias between the chips. As the temperature varies across the detector it causes small variations the resistance of each chip which, in turn, changes the proportion of the overall bias voltage applied to each chip.

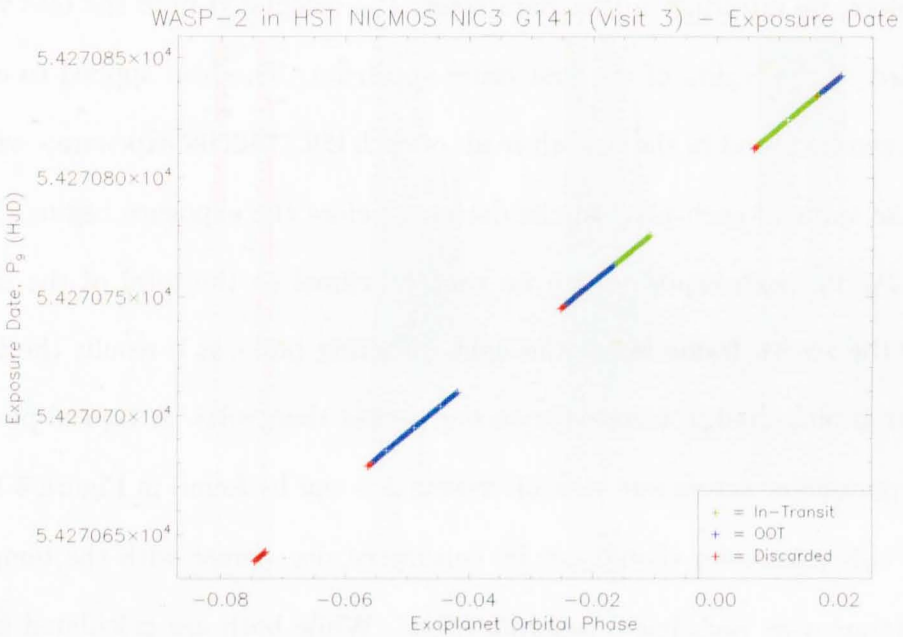


Figure 4.12: *A typical plot of the Exposure Time of each image in a single HST visit of observation, given by the Julian Date of each image. Points on the graph are coloured as in Figure 4.2.*

While Burke et al. (2010) address this problem by identifying the state of each exposure and normalizing the exposures in each state independently, their results indicate a dependency on the bias applied to the region of the chip containing the first order spectrum, rather than a dependency on the relative balance of the bias between the chips.

Therefore, we introduce a Bias parameter, P_{10}, calculated from the bias voltage applied to the region of the first order spectrum. The bias applied to each exposure can be found in the 'zeroth-read' of each `MULTIACCUM` exposure - which records the value of each pixel in the detector before the exposure begins. The value of P_{10} for each exposure can be simply defined as the total of the pixel values in the zeroth frame inside the light-gathering box - as a result, the Bias parameter is measured in counts-per-second rather than volts. A typical plot of the Bias parameter across one visit of observation can be found in Figure 4.13.

Note, this parameter should not be considered degenerate with the temperature parameter as each has a separate effect. While both are calculated from the bias voltage in the zeroth read of each `MULTIACCUM` exposure, the resulting temperature value is for the detector as a whole, while the Bias parameter applies only to the local region of the light-gathering box. As such, the Bias parameter can be considered the local parameter to the temperature's global measure.

4.2.4 Parameter Summary

For sake of consistency, the detrending parameters will hereafter be referred to by their P_x designation. However, to allow easy reference to the source of each

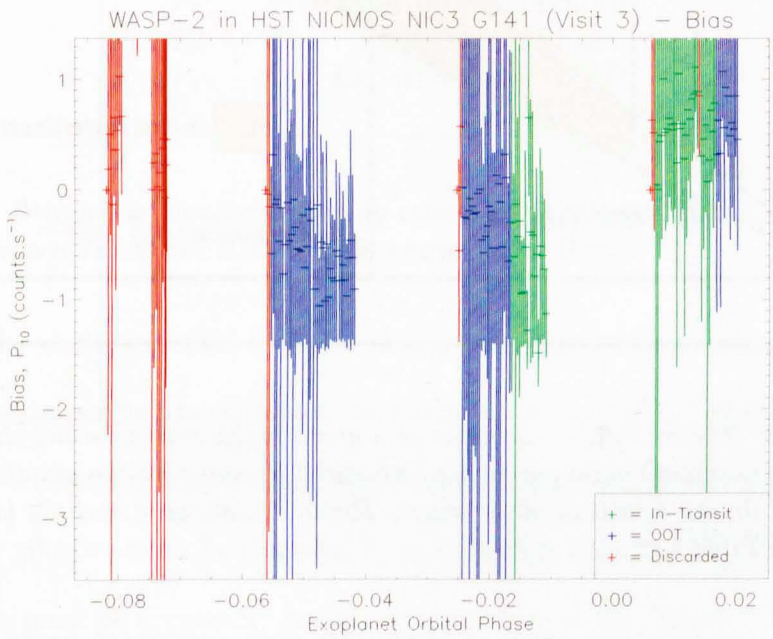


Figure 4.13: *A typical plot of the counts recorded as a result of the Bias at the location of the first order spectrum in the zeroth-frame of each exposure across a single HST visit of observation. Points on the graph are coloured as in Figure 4.2.*

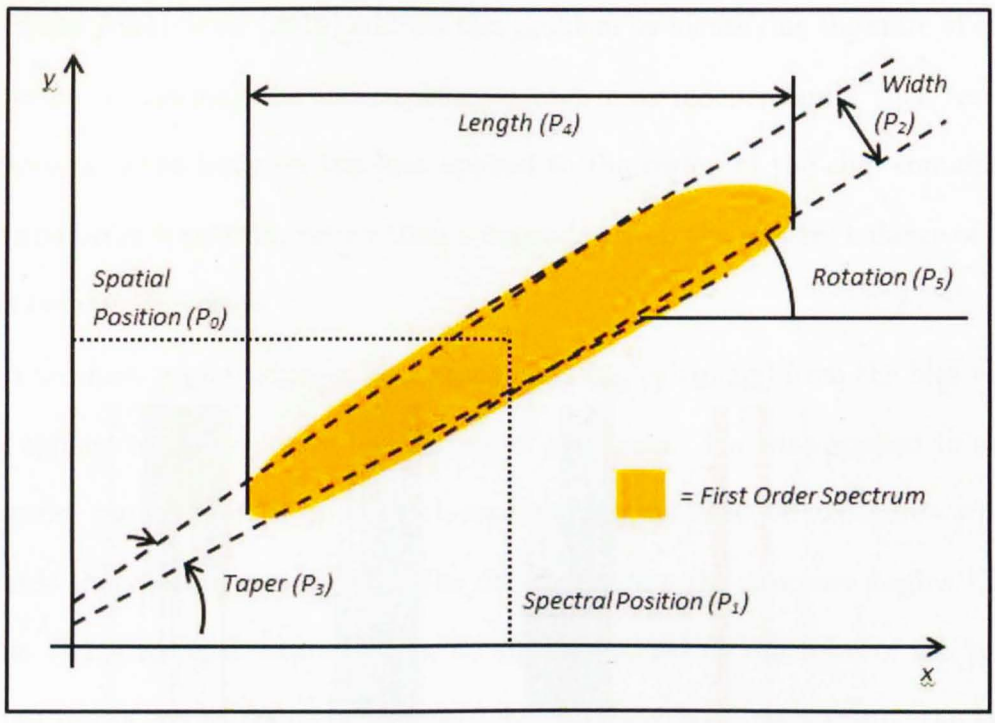


Figure 4.14: *The first order spectrum of a grism image (not to scale) shown in relation to the systematic noise parameters measured to constrain the pointing variations experienced during a visit of observation. The full meaning of each P_x label can also be found in Table 4.1.*

parameter, the following table (4.1) equates each parameter to its P_x number and Figure 4.14 demonstrates the quantities measured by each of the parameters arising from pointing variations.

4.3 Removal of Systematic Noise

Once we have properly parameterised potential sources of systematic noise, we must ensure that such noise is removed without influencing the transit itself. The impact of the systematics can be modelled simply by examining direct relations

Parameter Number	Parameter Name
P_0	FOS Spectral (y) Position
P_1	FOS Spatial (x) Position
P_2	FOS Width (in y)
P_3	FOS Taper
P_4	FOS Length (in x)
P_5	FOS Rotation
P_6	Detector Background at FOS
P_7	Detector Temperature
P_8	HST Orbital Phase
P_9	Exposure Date (HJD)
P_{10}	Bias

Table 4.1: *Detrending Parameter look-up table. (FOS = First Order Spectrum. HST = Hubble Space Telescope. HJD = Heliocentric Julian Day.)*

between the change in flux between images and the corresponding changes in the various detrending parameters. However, since the transit itself is a change in flux, in-transit exposures must be excluded from this process. Therefore, if systematic effects are to be removed from in-transit images, the effects of these systematics must be accurately modelled using the Out-Of-Transit (OOT) images alone.

4.3.1 Linear Regression

The simplest means of removing the systematic noise - or "detrending" the light curve - is to define a simple model to describe the impact of our measured systematic parameters on the measured flux of the light curve. If we assume that our ideal, systematic-free light curve would have no variation in flux across all Out Of Transit (OOT) images and assuming all of our systematic parameters are

independent, both of each other and of the actual light falling on the detector, we can state that:

$$C(i) = \overline{C}(i) + \sum_{j=1}^{q} P_j(i) Z_j \tag{4.12}$$

where $C(i)$ is the counts measured in an image, i; $\overline{C}(i)$ is the (as yet unknown) corrected counts of that exposure, with no systematic noise; $P_j(i)$ is an array of our parameters, j, for the image; q is the number of parameters (in our case, $q = 9$); and Z_j is an array of constants describing the correlation, if any, between each parameter and the recorded flux.

For all i and j both C and P are already known. Our ideal light curve is given by F. Therefore, in order to extract this light curve from the measured counts, we must solve the above equation to find Z. As addressed earlier, we are assuming that all OOT points in our ideal light curve are equal to 1. Therefore, we can calculate Z for the OOT images and apply the result to determine the correct F for the in-transit images.

The required solution can be easily acheived through a process of Linear Regression, as originally implemented by Gilliland (2005). However, the light curve resulting from this process is not completely detrended.

Since the detrending parameters (P_0 - P_8) have themselves been calculated from the very images we seek to detrend, we must expect a certain degree of systematic noise to also exist in the parameters themselves. One stark example can be found in Figure 4.15 which demonstrates a bimodal temperature parameter

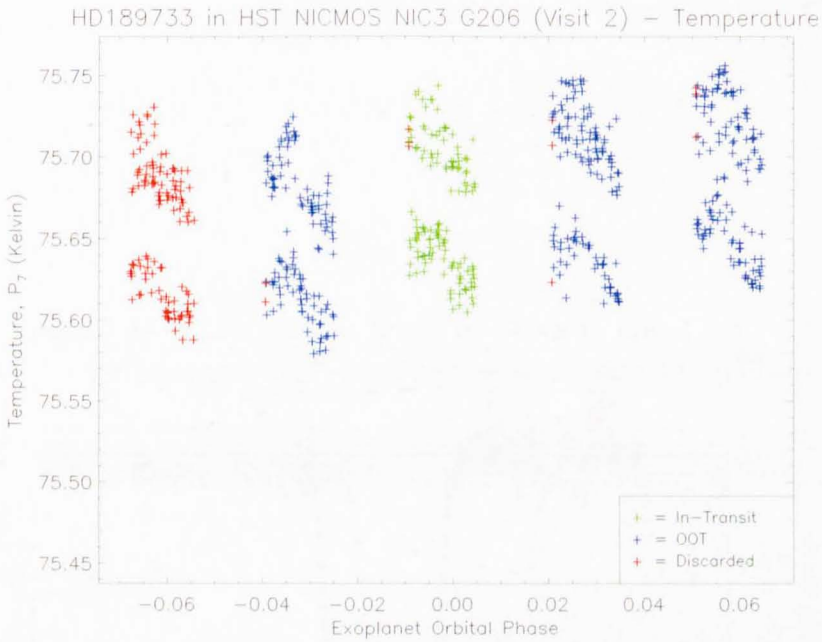


Figure 4.15: *The temperature across a single visit of HST observations in the same format as Figure 4.2. Two populations of exposures are evident in each orbit - resulting from the bias voltage switching between different states.*

- the result of systematic effects in the bias voltage for that visit. This means that the original assumption that the parameters were independent, both of each other and the recorded flux, is false.

In effect, Linear Regression performs a first-order detrending process and is capable of removing a large proportion of systematic noise from its subject light curve, as shown in Figures 4.17 and 4.16. However, with only a first-order approach, it is possible that some systematic effects may remain after the Linear Regression process.

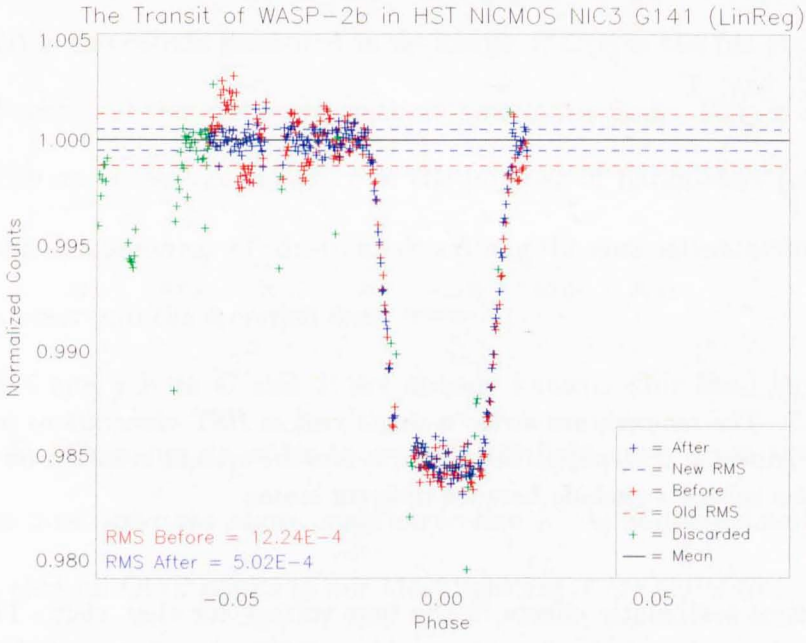


Figure 4.16: *A transit light curve shown before (red) and after (blue) the removal of systematic noise through a process of Linear Regression. Those points shown in green were identified as anomalous and removed from future calculations prior to the detrending process (as discribed in Chapter 3). The final regressed light curve alone can be found in Figure 4.17.*

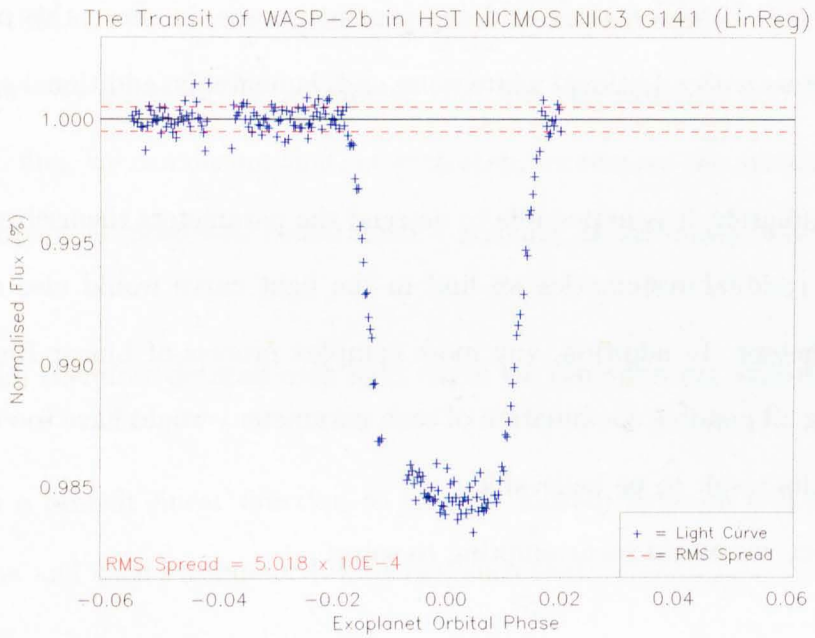


Figure 4.17: *A transit light curve after the removal of systematic effects by a process of Linear Regression. From the slant of the transit floor in particular, it is evident that some trace systematics likely remain.*

4.3.2 Preferential Decorrelation

If we are to ensure that all systematic effects are removed as effectively as possible, we must consider the shortcomings of the Linear Regression process. A perfect set of parameters would, of course, result in a perfect detrending process and a final light curve affected only by white noise and any real changes in flux resulting from the star itself. However, the calculated parameters are, in effect, this perfect set of parameters with additional white noise and, lamentably, additional systematic noise.

Unfortunately, it is impossible to detrend the parameters themselves because the same residual systematics we find in the light curve would also remain in the parameters. In addition, any more complex process of Linear Regression - containing all possible combination of each parameter - would have too many free terms for its result to be believable.

In effect, we would be attempting to solve:

$$C(i) = \overline{C}(i) + \sum_{j=1}^{q} P_j(i)Z_j + \sum_{j=1}^{q}\sum_{k=1}^{q} P_j(i)P_k(i)Z_{j,k} + \sum_{j=1}^{q}\sum_{k=1}^{q}\sum_{l=1}^{q} \cdots \quad (4.13)$$

This leaves us with the option of dealing with the sources of systematics iteratively, such that:

$$C_n(i) = C_{n-1}(i) - P_j(i)Z_{j,n} \quad (4.14)$$

where C_n is the counts in the current (nth) iteration; C_{n-1} is the counts from the

previous iteration; P_j is a selected detrending parameter; and $Z_{j,n}$ is a coefficient defining the impact of that parameter, unique to the current iteration.

After an infinite number of iterations, $C_\infty(i)$ will be equal to $\overline{C}(i)$ in Equation 4.3.2. However, arriving at this solution via an iterative route allows us to ensure that the solution is believable.

Note that in each iteration, the light curve is only detrended against a single parameter. By selecting the parameter that has the strongest correlation with the OOT flux, we can ensure that, at each step, we remove the systematic effect whose impact is most well constrained - arriving at the most believable light curve.

We can therefore detrend each light curve via the following algorithm:

1. Fit a best-fit linear function to the relationship between Out Of Transit Flux and each Parameter (P_0 to P_8), such that:

$$C(i_{OOT}) = P_j(i_{OOT})Z_j + r_j \tag{4.15}$$

   where $C(i_{OOT})$ is the OOT light curve; $P_j(i_{OOT})$ is the corresponding parameter set for a parameter, j; Z_j is the gradient of the best-fit line; and r_j is the constant term for this relationship.

2. Calculate a measure of correlation for each parameter with flux, such that:

$$\rho_j = \frac{\sum_{i=0}^{n-1}(C(i_{OOT}) - \overline{C_{OOT}})(P_j(i_{OOT}) - \overline{P_{j,OOT}})}{\sigma_{C,OOT}\sigma_{j,OOT}} \quad (4.16)$$

where ρ_j is the correlation coefficient for a parameter, j; $C(i_{OOT})$ is the counts recorded in an out-of-transit image, i_{OOT}; $\overline{C_{OOT}}$ is the average out-of-transit counts across all exposures, n; $P_j(i_{OOT})$ is the value of parameter j for those out-of-transit exposures; $\overline{P_{j,OOT}}$ is the average value of that parameter; $\sigma_{C,OOT}$ is the standard deviation of the out-of-transit counts; and $\sigma_{j,OOT}$ is the standard deviation of the out-of-transit parameter.

3. Select the parameter with the greatest value of ρ and remove its modelled linear trend from the whole light curve (both in-transit and out-of-transit), such that:

$$C'(i) = C(i) - P_j(i)Z_j - r_j + \overline{C}(i_{OOT}) \quad (4.17)$$

where $C(i)$ is the input light curve; $C'(i)$ is the light curve for the next iteration; $P_j(i)$ is the values of the selected parameter, j; Z_j and r_j are the coefficients of the best-fit line for that parameter; and $\overline{C}_{OOT}$ is the mean OOT Flux - typically already normalised to unity.

4. Calculate a measure of the difference between the new light curve and its predecessor, such that:

$$d^2 = \frac{\sum_{i=0}^{N-1}(C(i) - C'(i))^2}{N} \tag{4.18}$$

where d is the difference; C is the light curve; and C' is the new light curve, both for a series of N images.

5. Repeat steps 1-3, omitting the parameter with the maximum value of ρ, but reintroducing the parameter omitted during the previous iteration.

6. Continue iterations until the difference, d, between iterations falls below a threshold value defined in the config file (typically 1% of the Poisson limit).

   This means of terminating the iterations is preferable to simply waiting for the spread of the points to fall below a certain limit because, without an accurate model for the transit itself, the spread of the light curve can only be measured accurately for the out of transit points. As such, any threshold value based on the RMS spread of the light curve could potentially stop the iterations with systematic effects still present in the in-transit points.

   Alternatively, a maximum number of iterations can be specified by the user via the config file.

I call this procedure "Preferential Decorrelation", due to its selective, iterative approach to detrending. From Figure 4.18, it is evident that no discernable residual correlation remains after the application Preferential Decorrelation. In comparison to Linear Regression (see Figure 4.19) the Preferential approach produces a more reliable light curve with a superior RMS Spread.

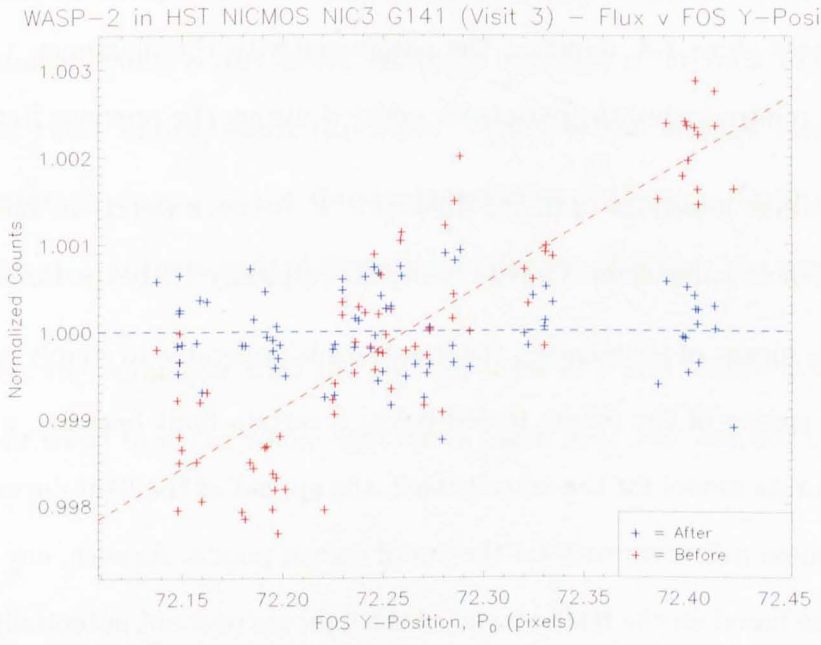


Figure 4.18: *A plot of Out-Of-Transit (OOT) Flux against Spatial (y) Position both before (red) and after (blue) a process of preferential decorrelation. The effect of the y-position systematic is evident in the gradient of the red dashed line. In contrast, after detrending (as demonstrated by the blue dashed line), that correlation has been removed.*

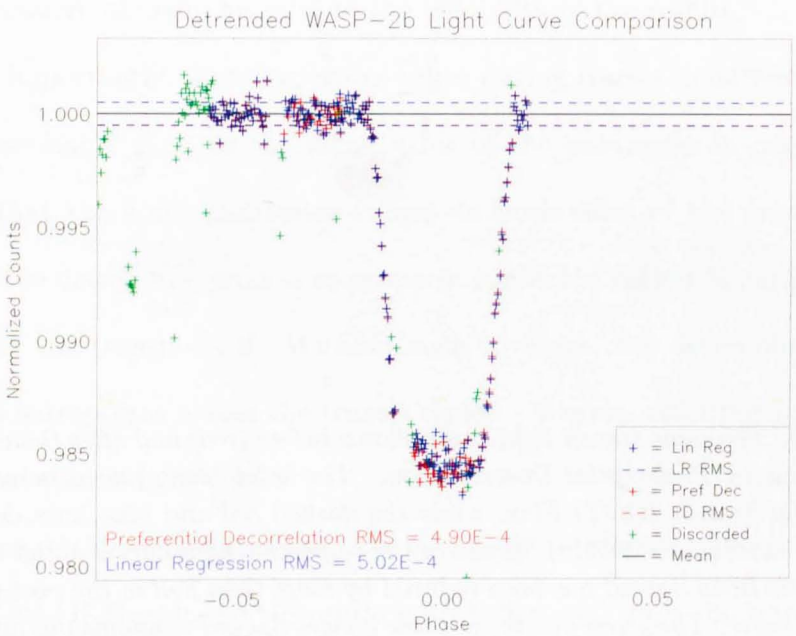


Figure 4.19: *The same transit light curve as found in Figures 4.16 and 4.17 with systematic effects removed by processes of Preferential Decorrelation (red) and Linear Regression (blue). Both curves demonstrate an improvement in RMS spread for the out-of-transit points, with the curve detrended by preferential decorrelation slightly superior. Most notably, the transit floor of the red curve does not exibit the same residual systematics as seen in the blue (linear regression) curve.*

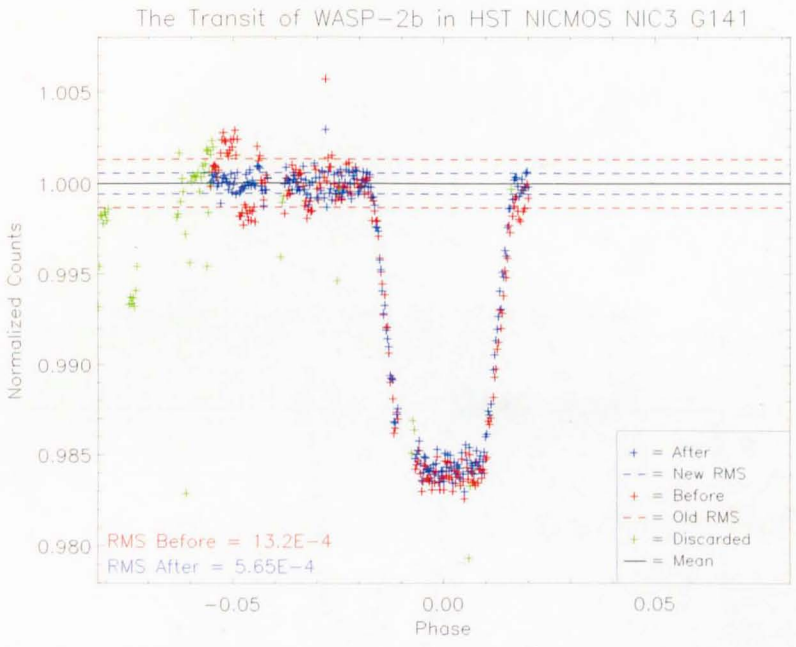


Figure 4.20: *The same transit light curve shown before (red) and after (blue) a process of detrending by Preferential Decorrelation. The solid black line demonstrates the mean Out-Of-Transit (OOT) Flux, while the dashed red and blue lines demonstrate the Root-Mean-Squared (RMS) Spread of the respective light curves about that mean. Note that the RMS Spread has been reduced by more than half in the post-detrending (blue) light curve. The green points are those images flagged as anomalous and removed from processing before the detrending process.*

A perhaps more effective means of demonstrating the effect of the detrending process can be found in Figure 4.20 which shows the detrended light curve overplotted on its systematic-beset predecessor. Further comparisons between the Preferential Decorrelation and Linear Regression techniques will follow in Chapter 5.

4.3.2.1 Caveats

It should be noted that either detrending process can only function properly with a comprehensive set of parameters with which to model the systematics. As such, there are various factors which can cause the detrending process to be ineffective or, on occasion, actually harmful to the reliability of the results.

Most importantly, those exposures taken during transit must have parameter values reasonably close to the mean value of the parameter in question. This ensures that there are parameter values on both sides of the transit values - allowing the detrending process to properly model the region of parameter-space containing the transit itself. Without such coverage, the detrending process is forced to extrapolate across the transit region - thereby reducing the reliability of the final transit light curve.

To ensure this does not occur, it is advisable to plan the observing strategy to have Out-Of-Transit (OOT) orbits both before and after the transit itself. If possible, it is also beneficial to time the observations to place a small number of OOT points in every orbit - either just before the ingress or just after the egress - to better account for parameters which have greater orbit-to-orbit variation.

In cases where good coverage is impossible, it is advisable not to detrend any in-transit points at all. In such cases, certain parameters can become degenerate out of transit - exhibiting the same patterns and frequency - while their in-transit values are still disparate. In such instances, the preferential decorrelation algorithm would have no means to determine which of the two degenerate parameters was responsible for an observed trend in the out-of-transit data and may select the

incorrect parameter. In such cases, the out-of-transit data would be detrended properly, but the in-transit data would likely become corrupted.

With no clear difference between such degenerate parameters, the resulting transit cannot be believed. In particular, the two temporal parameters (Orbital Phase and Visit Time) are guaranteed to be degenerate in visits with only one orbit containing OOT exposures.

Another factor prone to adversely affecting result is the magnitude of the parameters themselves. Since the majority of systematics arise from pointing variations, their correlation can be reliably modelled by a straight line - since we are only concerned with the variation in sensitivity between one pixel and the next. However, should such pointing variations become too large, light falling in one pixel at the start of a visit of observation may finish that visit falling two or more pixels away. In such cases, a straight line would not be a reliable means of modeling the correlation between the measured flux and the corresponding parameters.

Thankfully, pointing variations are almost exclusively on the sub-pixel scale, so this effect is rarely a significant factor. In fact, variations on this scale typically only occur in the first orbit of any visit of observation - a prime example of why this orbit should be removed from processing.

4.4 The Final Light Curve

An ideal process of Preferential Decorrelation is capable of reducing the noise in a light curve to $\sim 2\sigma_{Poisson}$ (roughly twice[1] the expected Poissonian white noise). However, this figure is dependent on a comprehensive set of parameters with good coverage. Exact levels of detrending, and comparison with previous results, will be discussed in the relevant Results chapters.

Typically, any successful detrending process will produce a final light curve ready to be modelled, as shown in Figure 4.21.

[1]Reaching a precision of $1\sigma_{Poisson}$ would be impossible in practice since the systematic noise becomes impossible to model accurately once it drops below the level of the white (random) noise. A perfect detrending could therefore be expected to leave a finished light curve with $\sim 1\sigma_{Poisson}$ of random noise and a remaining $\sim 1\sigma_{Poisson}$ of systematic noise. The measured spread of the light curve is likely to be greater still; due to fluctuations in stellar intensity and other real variations in the measured flux (aside from the transit itself).

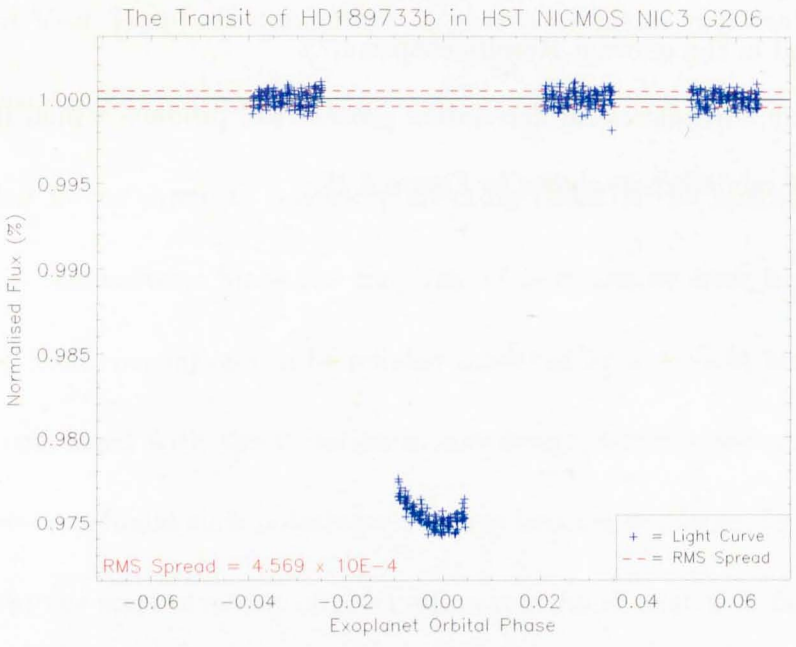


Figure 4.21: *A typical final, detrended transit light curve comprising two overplotted HST visits of observation - in this case, the transit of HD189733b. The data have been phase-folded about the known ephemeris of the exoplanet and out-of-transit points have been normalised to unity.*

Chapter 5

Methods 3 - Transit Modeling

5.1 Chapter Outline

The following chapter details the modeling of a processed transit light curve to derive the parameters of the planetary system which created it. An overview of the dependency of the transit on the various planetary and stellar parameters is followed by details on the generation of a model transit light curve from a given set of said parameters. A review of the basic χ^2-fitting technique precedes an in-depth description of the Markov-Chain Monte-Carlo algorithm applied to the light curves in this research. An introduction to the foundation of this technique - the Metropolis-Hastings algorithm - is followed by a review of the alterations made to improve the effectiveness and efficiency of the modeling process (specifically: the application of priors; automated step-size rescalling; and the rotation of parameter-space). The chapter concludes with a description of the derivation of the ideal parameter set from the MCMC results.

5.2 Planetary Parameters

With a complete, detrended transit light curve, we can now derive the parameters of the planet that created it. To this end, we create a model light curve whose parameters can be varied to create a best-fit to the observed transit.

The shape of the transit light curve is dependent on various properties of the observed extrasolar system. In particular, we are concerned with the radius of the planet (R_P) which can be determined as a ratio with the stellar radius (R_S) from the transit depth. The ingress and egress of the transit are dependent on the inclination angle (i) which determines the path of the planet across the disk of the star. As discussed in Chapter 1, the intensity of light across the star is dependent on the stellar limb darkening, which we characterise with to parameters (U_0 and U_1) via a quadratic law (Equation 1.3.2.3).

The transit is also dependent on the semi-major axis of the planet's orbit (a) and its orbital period (P). However, if we assume a fixed stellar mass (M_S) we need only include one of these parameters in our model, since the two are related by Kepler's third law:

$$P^2 = \frac{4\pi^2}{M_S G} a^3 \tag{5.1}$$

where P is the planet's orbital period; M_S is the stellar mass; G is the gravitional constant; and a is the planet's semi-major axis.

Finally, any variation in the transit midpoint (ΔT) - such as through errors in the ephemeris or through transit timing variations (see Chapter 1) - would

result in a non-central light curve (i.e. the planet's orbital phase at the transit midpoint would not equal 0 when calculated from the planet's ephemeris alone). Since our light curve model (see Section 5.2.1) assumes a transit symmetrical about this ephemeris-calculated midpoint, we must also model any offset in the actual transit midpoint (ΔT) to account for such effects.

Therefore, to properly parameterise the transitting planet, we must model:

1. The planetary radius: R_P.
2. The inclination angle of the planet's orbital plane: i.
3. The stellar radius: R_S.
4. The semi-major axis of the planet's orbit: a.
5. The parameters modeling the stellar limb-darkening: U_0 and U_1.
6. And an offset of the transit midpoint for each visit: ΔT_V, where V is the number of the visit in question.

For simplicity's sake, we will simply refer to these as a single parameter vector: $\mathbf{q_j}$, where j refers to the index of each parameter in the above list.

To determine the best values of these parameters, we must create a model transit light curve, Γ, to test against our observed light curve, L. Once such a best-fit model is found, the parameters of that model will be the best-fit solution for the properties of the observed planet and its orbit.

5.2.1 Generating a Model Light Curve

To generate our model light curve, we employ the analytic formula developed by Mandel and Agol (2002). This algorithm requires inputs of the ratio of radii between the planet and its parent star (R_P/R_S), two limb-darkening coefficients (U_0 and U_1) and the displacement of the centre of the disc of the planet from the disc of the star as a function of time, given as a fraction of the stellar radius ($D(t)/R_S$).

The value of $D(t)$ can be calculated for each point in the light curve from two components demonstrated in Figure 5.1, such that:

$$D(t) = \sqrt{b^2 + B^2(t)} \tag{5.2}$$

where $D(t)$ is the separation between the centres of the disk of the star and planet at a time t (given as a fraction of the exoplanet's orbital period, P, such that $-0.5 < t < 0.5$, as calculated in Equation 3.7); b is the impact parameter of the transit (the minimum separation between the centres of the star and planet); and $B(t)$ is the distance of the planet from the midpoint of its transit.

The impact parameter, b, can be calculated (see Figure 5.1) from the inclination angle (i) and the planet's semi-major axis (a), such that:

$$b = a\sin(i) \tag{5.3}$$

where b is the impact parameter; a is the semi-major axis of the planet's orbit; and i is the inclination angle of the orbit relative to our line of sight.

The time (t) is already converted into a fraction of the orbital phase and set to 0 at the transit midpoint. Therefore, if we define ϕ as the angular separation of the planet from the transit midpoint throughout one orbit (as shown in Figure 5.1) the value of $B(t)$ can be calculated as:

$$B(t) = a\sin(\phi(t)) = a\sin(2\pi t) \tag{5.4}$$

Therefore, we can express the seperation of the stellar and planetary disks as:

$$D(t) = a\sqrt{\sin^2(i) + \sin^2(2\pi t)} \tag{5.5}$$

The complex interplay between these inputs governs the shape of the transit light curve - as demonstrated in Figures 5.2, 5.3 and 5.4.

5.3 χ^2 Fitting

The simplest means of finding a best-fit solution for our transit light curve is to assess the χ^2 goodness-of-fit statistic of our model to the data (Press et al., 2007a). For any given model, the χ^2 statistic has the form:

$$\chi^2_j = \sum_{i=0}^{n-1} \left(\frac{(L(i) - \Gamma_j(i))^2}{\sigma(i)} \right) \tag{5.6}$$

where χ^2_j is the goodness-of-fit statistic for a set of model paramters, j; $L(i)$ is the value of the measured light curve from exposure i for a total of n exposures;

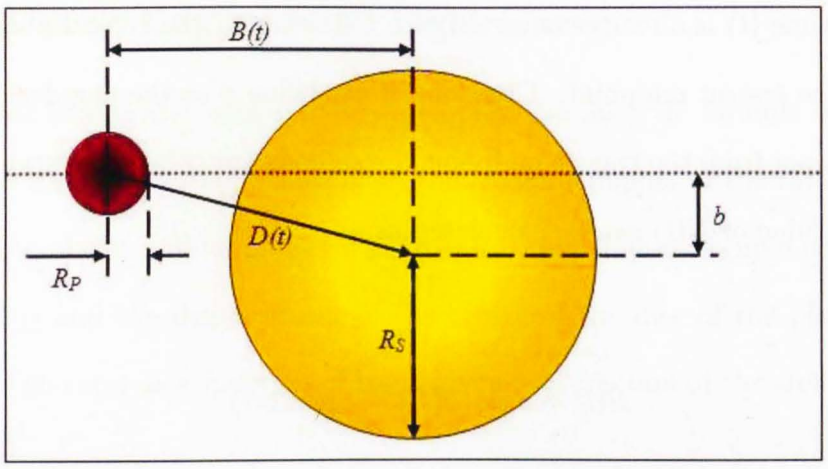
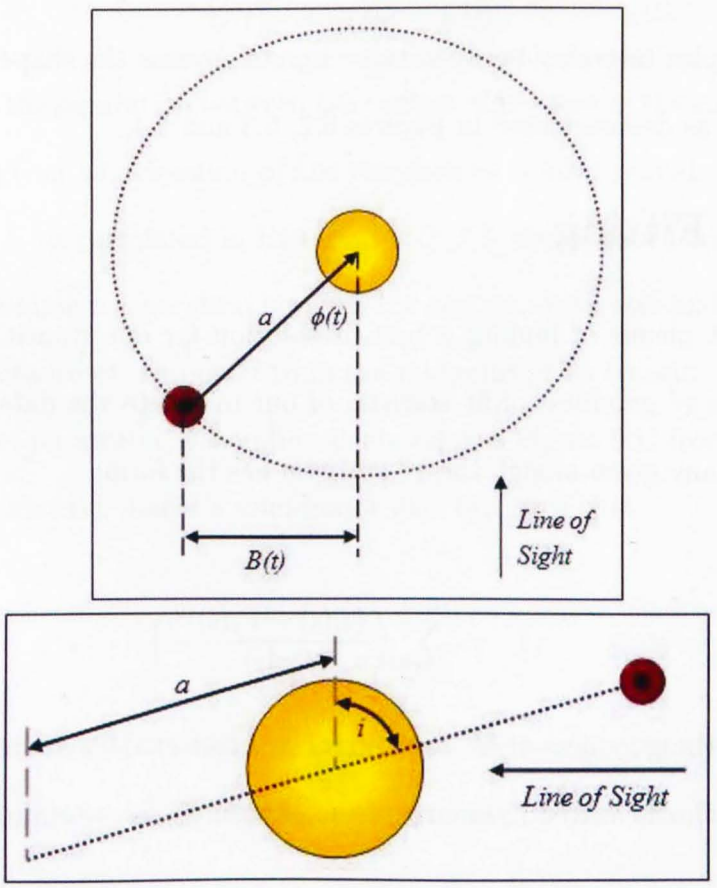


Figure 5.1: *The geometry of a transit event (above) demonstrating the components of $D(t)$ as the time-dependent orbital path ($B(t)$) and the time-independent impact parameter (b). The diagrams below demonstrate the calculation of $B(t)$ (top) from the planet's orbit and b (bottom) from the planet's inclination.*



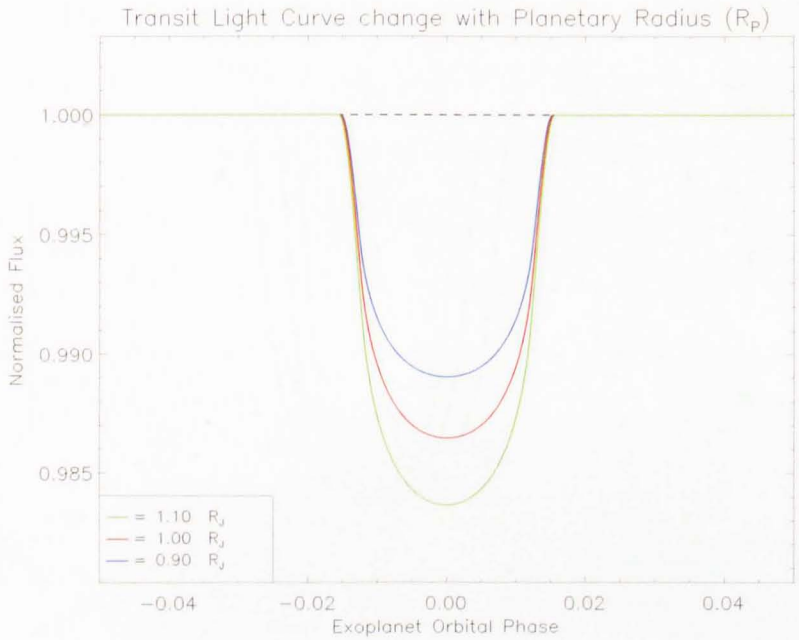
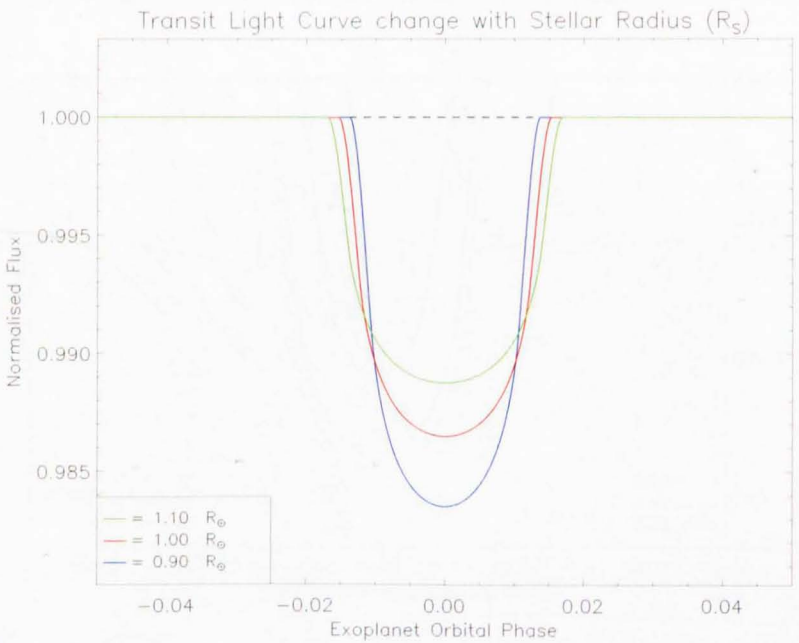


Figure 5.2: *Two model light curves generated by the Mandel & Agol algorithm. Each shows a standard model light curve (red), with two alternate light curves (blue and green) demonstrating the change in shape produced by the modification of one parameter. Above, the light curves show the transits produced by different Planetary Radii (R_P) and below, the effect of different Stellar Radii (R_S).*



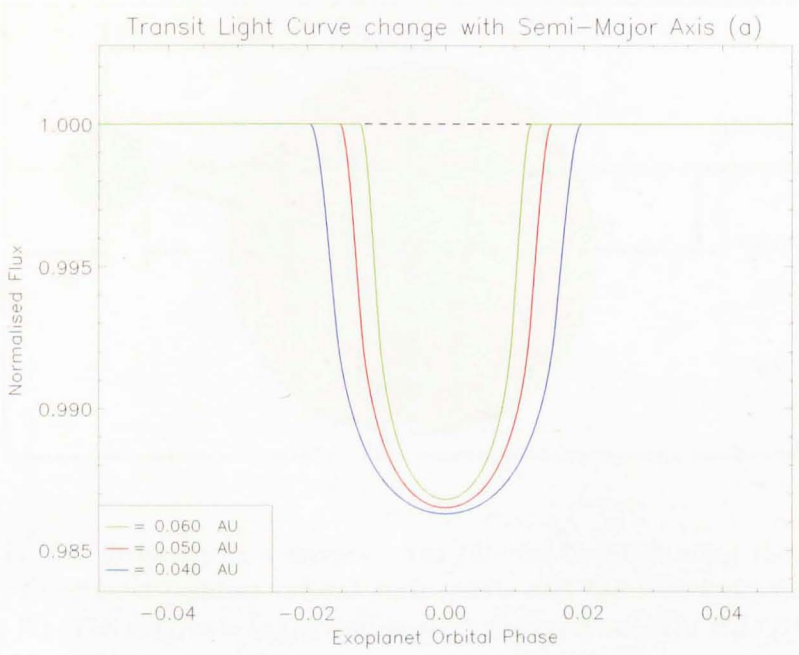
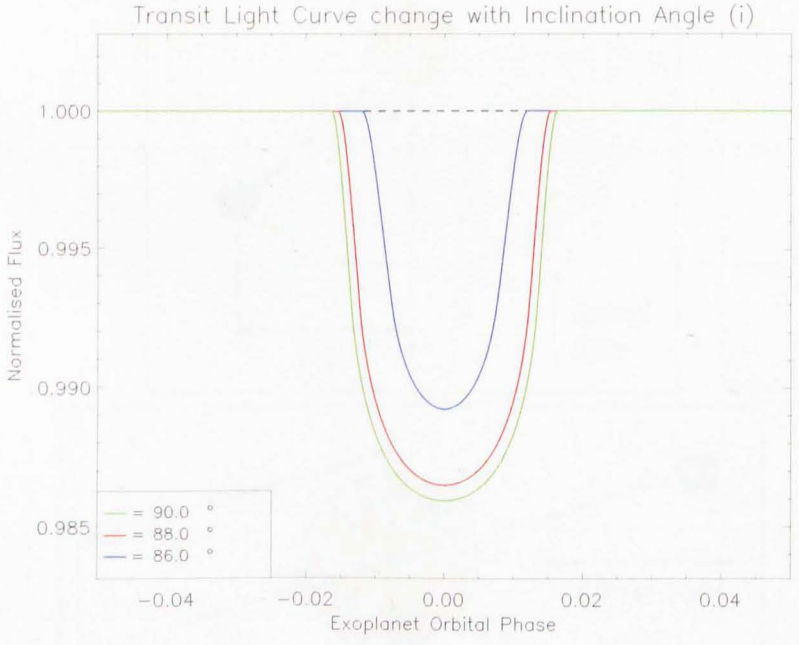


Figure 5.3: *The same standard model light curve as seen in Figure 5.2, employing the same colour scheme. Above, the light curves demonstrate the transits produced by different semi-major axes (a). Below, the light curves demonstrate different inclination angles (i).*



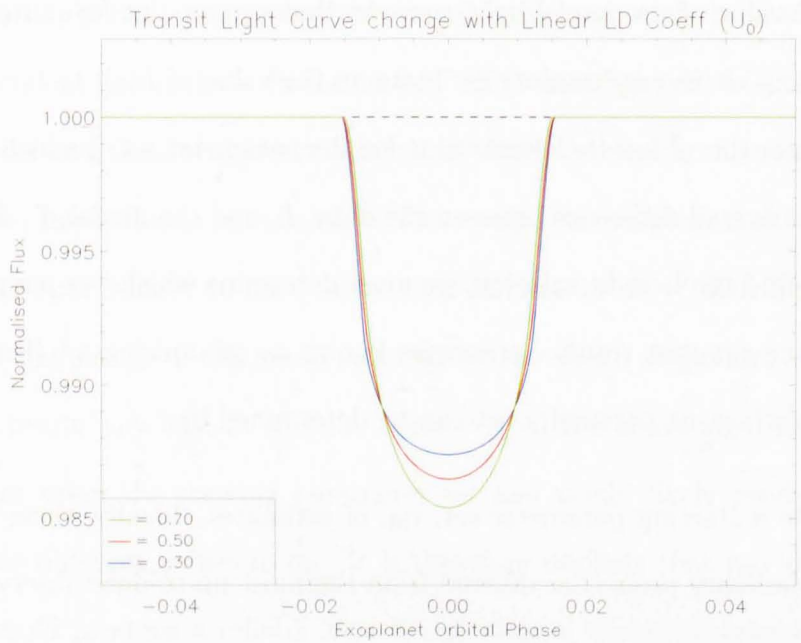
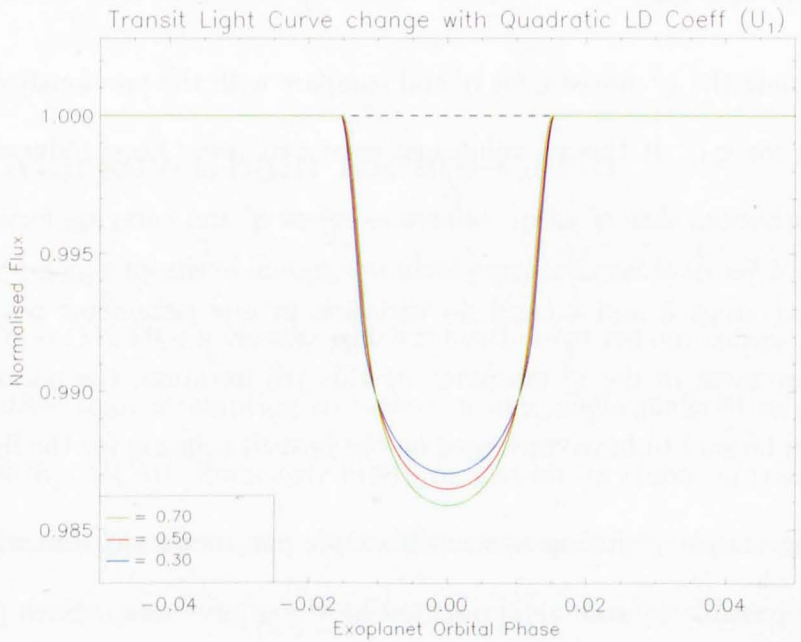


Figure 5.4: *The same standard model light curve as seen in Figure 5.2, employing the same colour scheme. Above, the light curves demonstrate the transits produced by different linear limb-darkening coefficients (U_0). Below, the light curves demonstrate different quadratic limb-darkening coefficients (U_1).*



$\Gamma_j(i)$ is the value of the model light curve at that same point for our parameter set, j; and $\sigma(i)$ is the uncertainty estimate on the value of $L(i)$.

By design, the χ^2 has its lowest value for the parameter set, j which produces the smallest overall difference between the data, L, and the model, Γ. Therefore, in order to find the best-fit solution, we must determine whichever parameter set results in the smallest result - otherwise known as minimising χ^2 (Press et al., 2007b). This best-fit parameter set can be determined by:

1. Create a starting parameter set, $\mathbf{q_0}$, of estimates. Ideally, these would be the planetary properties derived from the most up-to-date observations.

2. Calculate the χ^2 statistic for $\mathbf{q_0}$.

3. Perturb the parameters in $\mathbf{q_0}$ by some small amount, derived from the uncertainty estimates on the current planetary properties. Call this new parameter set $\mathbf{q'}$.

4. Calculate the χ^2 statistic for $\mathbf{q'}$ and compare with the previously calculated value for $\mathbf{q_0}$. If the χ^2 value has improved (i.e. been reduced) by the perturbation, save $\mathbf{q'}$ as $\mathbf{q_1}$, otherwise reject $\mathbf{q'}$ and carry $\mathbf{q_0}$ forward.

5. Repeat steps 3 and 4 until no variation in any parameter results in an improvement in the χ^2 statistic. At this jth iteration, the parameter set, $\mathbf{q_j}$ can be said to have converged on the best-fit solution for the light curve.

However, simple χ^2 fitting assumes a simple parameter space in which there is only one possible solution - represented by the χ^2 minimum. Such parameter

spaces can only truly exist for simple curves governed solely by a small number of independent parameters. In the case of our transit light curve, the quantity of parameters and their complex relationships which determine the overall shape of the transit mean this is not the case. The result is a complex, multi-dimensional parameter space, likely to have many local minima in its χ^2 distributions rather than the singular local minimum required by the above alogirthm.

As a result, any best-fit solution derived from the above algorithm would be dependent upon the starting parameter set and would likely produce different results for different values in $\mathbf{q_0}$. It is therefore unlikely that any process of χ^2 fitting would produce a reliable best-fit solution in this circumstance. It would certainly be unable to produce meaningful estimates of the uncertainty of each best-fit parameter.

To determine a reliable solution, we must instead identify the most likely set of parameters to produce the transit we observe - hereafter known as the 'ideal' solution.

5.4 Markov-Chain Monte-Carlo

The most reliable means of finding our ideal parameter set is to use Markov-Chain Monte-Carlo (MCMC) analysis - which robustly explores our complex parameter space, rather than attempting to locate on any single prefered solution (Press et al., 2007b). MCMC iteratively produces a series (or chain) of trials, each representing a new parameter set - or a different point in parameter space. In each iteration, a new parameter set is generated. This 'trial' is either accepted and

carried forward into the next iteration, or rejected, in which case the last accepted trial is carried forward instead.

In the simplest case - the "Metropolis-Hastings" alogirithm - MCMC works as follows:

1. Generate an initial light curve from a starting parameter set, q_0, which roughly estimates the shape and depth of the transit. For HST observations, we would expect previous studies to have already produced constraints on the planetary and orbital parameters ($\mathbf{q_0}$), along with accompanying uncertainty estimates (σ). These values provide a more than adequate starting point for the MCMC process.

2. Calculate the χ^2 goodness-of-fit statistic between this model light curve, Γ_0, and the observed transit, L.

3. Create a new parameter set by taking a random 'step' through parameter space. The step size can be calculated as a small percentage ($\sim$ 1%) of the uncertainty values, $\sigma(j)$ on each of the starting parameters, $q_0(j)$. Mathematically, for each parameter:

$$q_k(j) = q_{k-1}(j) + s\xi\sigma(j) \tag{5.7}$$

where $q_{k-1}(j)$ is the value of parameter, j, in the previous step (so, in the first iteration, $q_{k-1}(j) = q_0(j)$); $\mathbf{q_k}$ is the new value of that parameter, j, in the current step; $\sigma(j)$ is the uncertainty value associated with that parameter; s is the step-size of each jump through parameter space (typically set at a starting value of 0.01); and ξ is a random number selected from a uniform distribution between -1 and 1.

Note that the step size itself has no effect on the final solution. A small step size will still be free to explore the entirety of the available parameter-space, but will likely take a large number of iterations to do so. The 1% value employed here is simple an arbitrary starting value to ensure the initial iterations are controlled. The step size itself will be discussed further in Section 5.4.1.3.

4. Ensure the new light-curve is physically acceptable - i.e. that the planet is still smaller than the star; the limb darkening coefficients are between 0 and 1; the inclination angle is not greater than 90°, etc... If $\mathbf{q_i}$ is unreal, return to step 3 and generate a new parameter set.

5. Generate a new light-curve, Γ_k, from the modified parameter set, q_k.

6. As before, calculate the χ^2 statistic between this model, $\Gamma_k(l)$, and the data, $L(l)$ (where l is the index of each point in the light curve).

7. Determine whether the new parameter set should replace the old. In a simple 'bootstrapping' χ^2 minimisation algorithm, we would simply reject

any new parameter set with an inferior (greater) value of χ^2 while accepting any which produced an improvement (reduction) in χ^2. Such a technique would soon converge on a solution, but this would not necessarily be the 'ideal' solution; given the number of free parameters and the complexities of the parameter space involved (see Section 5.3).

Instead, we endeavour to explore parameter space and so must ensure that the chain does not simply converge on the nearest, local χ^2 minimum. To this end, any trial with an improved χ^2 is still automatically accepted, however, trials with inferior values of χ^2 may be accepted or rejected according to the probability:

$$\gamma = \frac{e^{-\chi_k^2/2}}{e^{-\chi_{k-1}^2/2}} \tag{5.8}$$

where γ is our test variable - the probability of acceptance; χ_{k-1}^2 is the goodness-of-fit statistic for our previous model light curve, Γ_{k-1} - created in the previous $(k-1)$ iteration from parameter set $q_{k-1}(j)$; and χ_k^2 is the goodness-of-fit statistic for the new model light curve, Γ_k - created in the current (kth) iteration, from parameter set $q_k(j)$.

The value of γ can then be used to test the new trial against a random number chosen from a uniform distribution between 0 and 1. If the random number is less than γ, the trial is accepted, otherwise it is rejected. Note

that, for an improvement in χ^2, γ will always have a value greater than 1 and so the trial will always be accepted. In contrast, the greater the detriment to χ^2, the smaller the chance that the new trial will be accepted. The fraction of accepted trials is dependent on both the shape of parameter space and the step-size between trials, this will be discussed further in Section 5.4.1.

8. If the trial is accepted, return to step 3 and carry forward $\mathbf{q_k}$ to generate a new trial. Otherwise, discard $\mathbf{q_k}$, return to step 3 and generate the next trial from the existing parameter set, $\mathbf{q_{k-1}}$.

9. Continue until enough iterations have been performed to have comprehensively explored parameter space - dependent on the number of parameters, the degree of degeneracy between the parameters and the complexity of parameter space.

   In general, the more iterations performed; the more precise the result, but there comes a point where further iterations will have negligible impact. The exact determination of this point, and the typical number of trials required to guarantee precise results, will be discussed further in Section 5.5

Once complete, the MCMC procedure will have produced a large number of accepted trials, accurately presenting the probability-density of the parameter space it has explored. Our ideal solution will therefore lie where the probability density is greatest - equivalently, where the greatest percentage of trials were accepted (Press et al., 2007b).

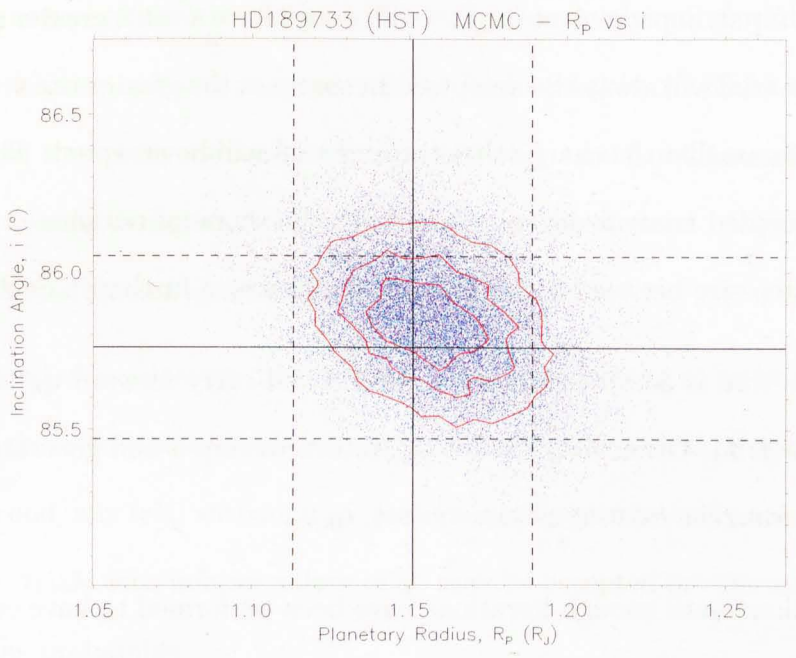


Figure 5.5: *A sample parameter-distribution plot showing a complete chain of accepted MCMC trials. Each blue point represents an accepted trial in parameter space between Planetary Radius (R_P, on the x-axis) and Inclination Angle (i, on the y-axis). The region with maximum density represents the region where trials were most likely to be accepted and therefore the most likely solution for both R_P and i. The probability distribution between these parameters is more readily visible in the overplotted contours, shown in red. The solid black lines represent the starting estimates on each parameter and the dashed black lines show the corresponding uncertainties to those estimates.*

The simplest means of displaying such probability densities is to plot two sets of accepted parameters against each other. This, in effect, produces a probability-surface which can be shown as a contour map (see Figure 5.5) with the peak of the surface representing the ideal solution for both parameters. The process by which this probability density is converted into an ideal parameter set will be discussed in Section 5.5.

5.4.1 Improving Efficiency

To ensure that the MCMC process produces an accurate map of parameter-space, a large number of iterations must be performed. If the final chain of accepted trials is too short, it will be unlikely to properly represent the parameter-space it has explored. However, it is unfeasible to perform an infinite number of trials to ensure perfect coverage. Therefore, it is imperative that the MCMC procedure be performed as efficiently as possible to conduct the largest number of trials possible in whatever time is available. To this end, several additional measures have been applied to ensure as efficient an exploration of parameter-space as possible.

5.4.1.1 Correlation Length

One means of quantifying the efficiency of the MCMC algorithm is to measure the correlation-length of the chain. The correlation length derives from the product-moment correlation between any one trial and each other step in the chain (Press et al., 2007a). In general, for any two distributions, the product-moment correlation coefficient between each has the form:

$$\kappa = \frac{\sum_k (U(k) - \overline{U})(V(k) - \overline{(V)})}{\sqrt{\sum_k (U(k) - \overline{U})^2}\sqrt{\sum_k (V(k) - \overline{V})^2}} \tag{5.9}$$

where κ is the correlation coefficient; $U(k)$ and $V(k)$ are two distributions with shared index, k; and $\overline{U}$ and $\overline{V}$ are the means of those distributions.

In effect, the correlation coefficient allows us to determine the number of steps taken between any two unique, independent trials. A large number of steps im-

plies the chain is taking too many steps to move across parameter-space, whereas a comparatively small number tells us the chain is more efficiently exploring the possible solutions to the transit fit.

The correlation coefficient of each trial can be easily calculated by comparing the residuals of each model light curve to its fellows. As such:

$$R(l) = L(l) - \Gamma(l) \tag{5.10}$$

$$\kappa_{j,k} = \frac{\sum_{l=0}^{n-1}(R_j(l) - \overline{R}_j)(R_k(l) - \overline{R}_k)}{\sqrt{\sum_{l=0}^{n-1}(R_j(l) - \overline{R}_j)^2}\sqrt{\sum_{l=0}^{n-1}(R_k(l) - \overline{R}_k)^2}} \tag{5.11}$$

where $\kappa_{j,k}$ is the correlation coefficient between trials j and k; $R_j(l)$ and $R_k(l)$ are the residuals for those trials (the measured light curve, L, minus the model light curve, Γ) at point, l in the light curve; and $\overline{R}_j$ and $\overline{R}_k$ are the mean values for those sets of residuals.

Note that, when $j = k$, the correlation coefficient will be 1; indicating that the trials are perfectly correlated. As we increase the number of trials separating j and k, the MCMC chain moves to a different part of parameter space and so begins creating model light curves of notably different shape. Once enough trials have been conducted, the chain will have lost all 'memory' of its previous shape and the correlation coefficient will have fallen to a reasonably constant

background level. The number of trials required for the chain to reach this point is known as the correlation length of the chain.

To measure the correlation length as accurately as possible, the value of κ is averaged over a large number of trials - for the same separation between j and k. This produces a correlation curve like those shown in Figure 5.6 which demonstrates an almost exponential decay from its starting position at 1 (as would be expected, since at iteration 0 the chain is being compared with itself, producing a correlation of 1) down to the background level at 0. This background level is the average correlation as $(k-j) \rightarrow \infty$ and, in effect, represents the state of minimum correlation between any two trials. For simplicity's sake, the curve is renormalised to set this value to zero.

A standard MCMC chain on a HST NICMOS quality light curve, resulting from the algorithm described in Section 5.4, can be expected to have a correlation-length between 50 and 100 trials (dependent on the number of in-transit points, the RMS spread of the data and the accuracy of the initial uncertainty measurements, σ_j - see Chapters 7 to 9). Ideally, we wish to engineer our exploration of parameter space so as to make the correlation length as small as possible. To this end, we apply the following modifications to the standard Metropolis-Hastings algorithm:

5.4.1.2 Priors

One source of inefficiency in the MCMC algorithm is the infinite boundaries on parameter space. Since every trial has a chance of being accepted, it is not

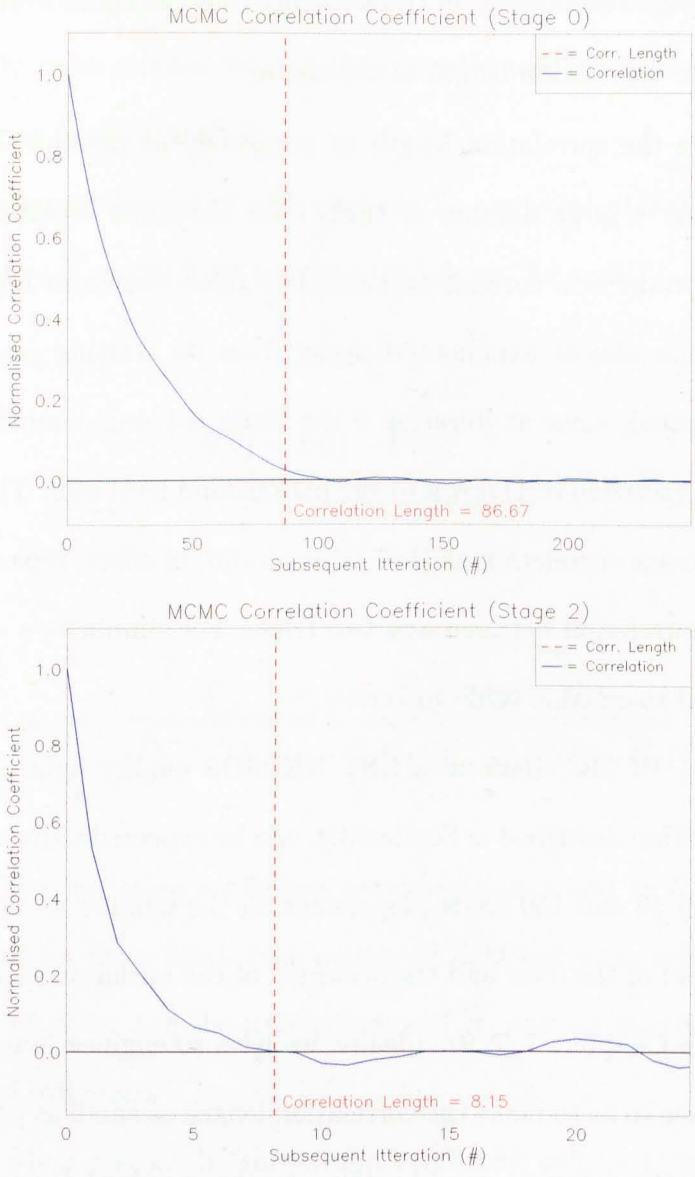


Figure 5.6: *The correlation curves (blue) for an inefficient (top) and efficient (bottom) exploration of parameter space. The dashed red line in each demonstrates the Correlation Length derived from each curve. Both curves have been normalized so that the value of the correlation coefficient at* ∞ *is equal to* 0 *(this would not normally be the case since the majority of out-of-transit points in the light curve will remain unchange - and therefore perfectly correlated - in every MCMC trial).*

impossible for the chain to wander very far from its ideal solution - and, indeed, the more trials that are conducted, the more likely this will be to occur.

While these trials still provide information about the shape of parameter space, they add very little to the accuracy of our final parameter values. With every step away from the ideal solution, the probability of acceptance diminishes while the equivalent region of parameter-space increases. As a result, any given area far from the ideal solution will contain such a small number of trials as to be statistically insignificant. Allowing the chain to explore this far therefore expends processing time for little-to-no gain.

Furthermore, allowing the chain to wander too far from its ideal solution may run the risk of finding other local minima with unphysical solutions - but which, nevertheless, result in light curves acceptable to the MCMC process. While such unphysical values can be easily identified and removed, much processing time can be wasted as the chain explores such unphysical local-minima.

Another similar problem occurs in the degeneracy between some of the parameters. In particular, since we are dealing with several measures of distance (R_P, R_S and a), scaling up each by the same percentage produces an ostensibly identical light-curve, but with potentially absurdly large or absurdly small values. Such degeneracies are unavoidable without fixing one of the parameters (though their effects can be managed) and typically result in deformation of the resulting probability distribution (see Figure 5.7).

For the most part, this can be easily managed by sanity-checking each trial as its parameter-set is generated. This step can automatically discard unrealistic

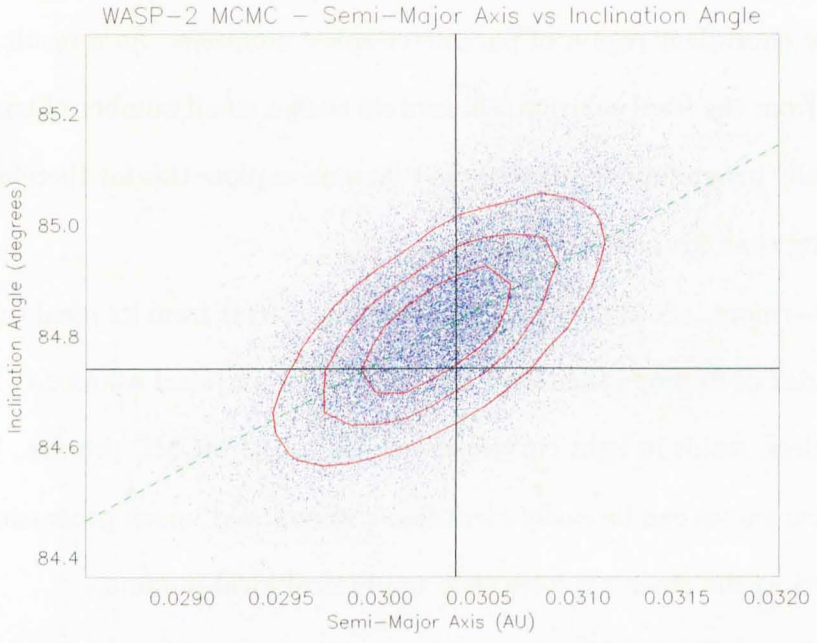


Figure 5.7: *An MCMC parameter distribution, as seen in Figure 5.5. However, unlike Figure 5.5, the relationship between Inclination Angle (i) and Semi-Major Axis (a) demonstrates clear degeneracy in the non-circularity of the distribution - modelled by the overplotted green, dashed line. The black lines denote the starting estimates of each parameter.*

values - such as Planetary Radii (R_P) greater than Stellar Radii (R_S); Inclination Angles (i) so great as to result in no transit; or Limb-Darkening Coefficients (U_0 and U_1) outside the range 0-1. However, this cannot prevent more subtle, but equally unphysical, parameter sets from forming part of the chain. To this end, we introduce priors (Press et al., 2007c).

Priors are additional probability weightings applied to the MCMC test variable, γ, (see Equation 7). They are a key aspect of all such Baysian statistics, allowing a probability distribution to be applied to a data set before the properties of that data are known. In effect, they are pre-defined controls on parameter-space, defined from the initial conditions of the MCMC trial.

For any HST transit data set, the parameters defining the shape of the transit will have been previously measured by other studies - providing a reliable starting point for our MCMC chain. The uncertainties on these values provide limits on the range of parameter space we might realistically wish to explore. We can therefore utilise these uncertainties to create an additional probability weighting for each accepted trial based on its displacement from our initial estimate. Typically, for simplicity's sake, we apply a simple Gaussian, such that:

$$\Omega_k(j) \quad = \quad e^{-\left(\frac{q_0(j)-q_k(j)}{s\sigma(j)}\right)^2} \tag{5.12}$$

where $\Omega_k(j)$ is the prior for the value $q_k(j)$ of a parameter, j, at MCMC iteration, k; $q_0(j)$ is the initial estimate of that parameter; $\sigma(j)$ is the initial uncertainty

on that parameter; and s is a weighting on the stringency of the prior (controlled by the `Config` file).

The overall prior, $\overline{\Omega}$ can then be calculated by combining the priors from each parameter being varied as part of the MCMC simulation. To ensure that the prior remains low if any one parameter is too far from the region of parameter space we wish to explore, the individual priors are combined as follows:

$$\overline{\Omega}_k = (\Omega_k(0)\Omega_k(1)\Omega_k(2)...\Omega_k(n-1))^{1/n} \tag{5.13}$$

where $\overline{\Omega}_k$ is the overall prior for the parameter set of iteration, k; $\Omega_k(j)$ are the individual prior values calculated for each parameter j via Equation 5.4.1.2; and n is the total number of parameters in j.

This combined prior ($\overline{\Omega}$) can then be applied to the test-variable, γ, to directly weight the probability of acceptance. The test-variable is therefore calculated as:

$$\gamma = \frac{\overline{\Omega}_k}{\overline{\Omega}_{k-1}} \frac{e^{-\chi_k^2/2}}{e^{-\chi_{k-1}^2/2}} \tag{5.14}$$

where $\overline{\Omega}_k$ is the value of the prior for the parameter k calculated for the parameter set of the current trial; $\overline{\Omega}_{k-1}$ is the value of the prior for that same parameter calculated for the previous trial; and where all other symbols hold the same meaning as in Equation 7.

Typically, the ratio of priors is small compared to the ratio of the χ^2 functions. However, as the MCMC chain moves further from the ideal solution, the difference

in χ^2 produced by a jump towards the ideal solution will result in a smaller improvement in χ^2. This reduces the effect of the ratio of χ^2 exponentials on γ and, instead, the ratio of priors becomes the dominant term; making it more likely that the MCMC chain will return to an area of parameter space closer to the ideal solution where the effect of the priors becomes minimal once more.

In some cases, a simple Gaussian prior is insufficient or inappropriate to properly control the explored region of parameter space. In particular, the limb darkening coefficients (U_0 and U_1) are unlikely to have reliable starting estimates or uncertainties from which to calculate such a prior. Therefore, for U_0 and U_1 we simply construct a uniform prior (with value 1) between 0.1 and 0.9, which then falls to zero at 0 and 1.

In effect, the priors apply an even-greater reduction to the chance of acceptance of those trials which fall far from the initial estimates - thereby reducing the percentage of trials which fall so far from the ideal solution as to be useless. By restricting, but not limiting, the explored range of parameter-space to that already constrained by previous observations, we also greatly reduce the potential frequency of unphysical local-minima and restrict the movement of the trials along degenerate parameter-relations. However, should previous results prove false or inaccurate, the procedure will still allow the MCMC chain to locate on a more preferable location.

The application of priors means that the resulting probability distributions are, in fact, the combination of the actual probability distribution and the prior itself. However, since the priors are pre-determined, their effects can be removed

before the ideal solution is calculated (see Section 5.5).

5.4.1.3 Rescaling the Step-Size

Another factor influencing the efficiency of the MCMC algorithm is the step size implemented for each parameter. In cases where the step-size is too small, the chain will take many iterations to move from one region of parameter space to the next. Conversely, if the step-size is too large, many trials will jump too far from the ideal solution to stand a realistic chance of being accepted.

An acceptance-rate of 25% produces the most efficient exploration of parameter-space (Press et al., 2007b) - allowing the chain to move quickly around the locale of the ideal solution while not moving so far as to have too many trials rejected. We can therefore define our optimum step-size as that which results in a 25% acceptance rate across all parameters.

The simplest means of arriving at this optimum is to introduce a weighting variable to the calculation (see Equation 3) of each new trial. By adjusting these weightings accordingly, the step-size can then be scaled to achieve the desired 25% acceptance rate. However, a simple global acceptance of 25% does not necessarily translate into the same acceptance for each parameter. It is therefore paramount to ensure a uniform acceptance rate across all parameters while also rescaling the parameter weightings to acheive the required global acceptance.

Globally, the parameter weightings can be rescaled with ease. Firstly, overall acceptance can be calculated as:

$$A(\alpha : \beta) = \frac{\Xi(\beta) - \Xi(\alpha)}{\beta - \alpha} \tag{5.15}$$

where $A(\alpha : \beta)$ is the acceptance rate for the MCMC trials conducted between iteration α and iteration β; and Ξ is the length of the MCMC chain (i.e. the number of trials accepted thus far) at those two iterations.

To ensure an accurate measure, a large number of trials (of order 10^4) are used in this calculation. So, for iteration k, $\beta = k - 1$ and $\alpha = k - 10^4$. From this acceptance rate, we can scale the step-size weightings to approach the desired acceptance rate, such that:

$$\omega'(j) = \omega(j)\frac{A(\alpha : \beta)}{A_0} \tag{5.16}$$

where $\omega'(j)$ is the improved weighting on the step size of parameter j; $\omega(j)$ is the existing weighting on that parameter; $A(\alpha : \beta)$ is the current global acceptance rate (as calculated in Equation 5.4.1.3); and A_0 is the desired acceptance rate - typically 25%.

These weightings can then be rescaled every 10^4 iterations - in line with the number of trials used to calculate the acceptance rate. If the value of A is too high, the modifier on the weightings (A/A_0) becomes greater than 1, thereby increasing the weighting, increasing the steps size and so causing the chain to

jump further from the ideal solution; thereby increasing the likelyhood that trials will be rejected and so decreasing the acceptance rate for the next recalculation. Conversely, if the acceptance rate is too low, the weightings are scaled down, thereby reducing the step-size and decreasing the number of rejected trials. Once the acceptance rate reaches the desired value, or thereabouts, the changes to the weightings will become negligible and the rescaling process can be said to have reached equilibrium.

However, even when in equilibrium, a global acceptance rate of 25% may, for example, represent a 50% acceptance rate for one parameter, balanced by a 12.5% acceptance rate for another. Therefore, we must also rescale the step-size weightings individually by parameter if we are to guarantee an efficient exploration of parameter-space.

Fortunately, individual rescaling does not require an individual measure of acceptance for each parameter (which would be impossible in the Metropolis-Hastings algorithm where all parameters vary concurrently). If the global acceptance rate is held at 25%, we need only ensure that the individual expectation of acceptance is the same for an equivalent step in each parameter to guarantee a uniform acceptance rate.

The expected acceptance for each parameter can be derived by recording the random numbers, ξ_j, applied to each step for each parameter in each iteration. Comparing the comparative size of each step (i.e. the magnitude of ξ_j) to the resultant global acceptance produces an acceptance curve, as shown in Figure 5.8.

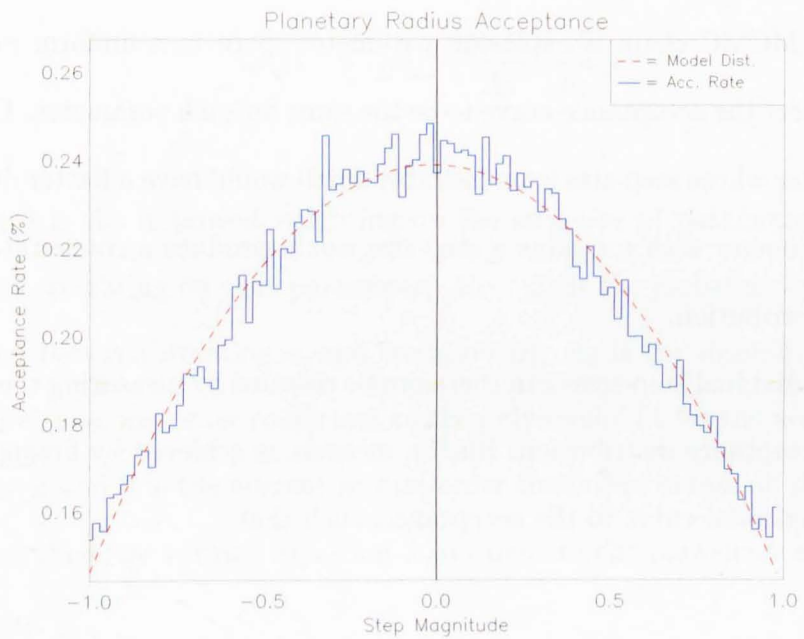


Figure 5.8: *A histogram (blue) of acceptance rate, A, binned by the magnitude of the random numbers, ξ, applied to calculate each step for the Planetary Radius parameter, R_P, across 10^4 MCMC iterations. The dashed red line shows the second-order polynomial fit to the histogram, which is later used to calculate the comparative acceptance-rate for this parameter.*

As would be expected, a larger jump (i.e. when $|\xi|$ is closer to 1) should be less likely to result in an accepted trial; as large jumps are more likely to move the chain away from the ideal solution. This is demonstrated by the curvature of the acceptance curve; peaking for very small jumps (which produce small changes in χ^2 and are therefore more likely to be accepted) and at its lowest for large jumps.

If our MCMC chain is exploring parameter space in a uniform manner, we would expect the acceptance-curve to be the same for each parameter. Conversely, a parameter whose step-size is undesirably small would have a flatter distribution and a parameter with too large a step-size would produce a comparatively more peaked distribution.

The individual step-sizes can therefore be rescaled by measuring the curvature of each acceptance distribution. Such a measure is achieved by fitting a second-order polynomial curve to the acceptance, such that:

$$A(\xi_j) = \mu_j \xi_j^2 + \nu_j \tag{5.17}$$

where $A(\xi_j)$ is the acceptance for a random number of magnitude ξ_j governing the step-size of a parameter j; μ_j is the second-order coefficient of the polynomial fit for this parameter; and ν_j is the constant-term of that fit.

Note that the linear term has been omitted from the best-fit curve. This is because we assume parameter space to be roughly symmetric and so would expect our acceptance curve to be identical for both positive and negative values of ξ.

While this is an imperfect assumption, its effects can be accounted for by other means (see Section 5.4.1.4).

Applying these individual measures of acceptance, our parameter weightings become:

$$\omega\prime(j) = \omega(j)\frac{A(\alpha:\beta)}{A_0}\frac{\overline{\mu}}{\mu_j} \tag{5.18}$$

where $\omega'(j)$ is the improved weighting on the step size of parameter j; $\omega(j)$ is the current weighting on that parameter; $A(\alpha:\beta)$ is the global acceptance rate (calculated between iteration α and iteration β); A_0 is the desired acceptance rate; μ_j is the second-order coefficient to the polynomial fit for the acceptance of parameter j; and $\overline{\mu}$ is the average second-order coefficient across all parameters.

We can therefore rewrite Equation 3 to calculate the parameter set for each new trial as:

$$q_k(j) = q_{k-1}(j) + 0.01\omega(j)\xi_j\sigma(j) \tag{5.19}$$

where $q_{k-1}(j)$ is the value of the parameter j from the previous $(k-1)$ iteration; $q_k(j)$ is the new value of that parameter for the current (k) iteration; $\omega(j)$ is the weighting on the step-size of this parameter; $\sigma(j)$ is the existing uncertainty estimate on this parameter; and ξ_j is a uniformly-distributed random number between -1 and 1.

Upon correct application, rescaling the step-size of the MCMC chain can result in a drastic improvement of the correlation length. While the measure is somewhat dependent on the efficiency of the initial step-size for each parameter, this step is typically capable of reducing the correlation length to less than 10% of its original value.

5.4.1.4 Rotating Parameter-Space

As noted in section 5.4.1.3 parameter-space is most efficiently explored when it is spherically symmetric - i.e. when the only factor effecting the shape of parameter space around the chain's current location is the ideal solution itself, and not degeneracies or false local minima.

However, in reality, parameter-space is likely to have a somewhat complex form and may be skewed (in particular by parameter degeneracies) into an eliptical or lenticular shape. While the degeneracies themselves cannot be removed, this does not mean that the situation is impossible to correct. Since the parameter-space itself cannot be changed, we can instead change the manner in which it is explored.

In the relatively simple Metropolis-Hastings algorithm, parameter-space is explored in a spherical manner - using random numbers to jump along a vector to a new location. This approach gives each step an equal chance of jumping the same distance in every direction. If our parameters are completely independent, and if we were to ignore the χ^2 minimum of our ideal solution, this would translate to each jump of equal length having an equal chance of acceptance, regardless of

direction. In reality, however, parameter degeneracies in particular (see Figure 5.7) would skew this probability; making jumps which move along the degeneracy more likely to be accepted and reducing the acceptance rate for jumps perpendicular to that relationship.

To efficiently explore parameter-space, we need to remove this dependence on direction and, since we cannot change parameter space itself (without risk of changing the result of the MCMC process), we must instead change the probability distribution of each jump through that parameter space.

In effect, for each new trial, we must modify the region available for each step to mirror the overall shape of parameter space. After a significant number of trials, this shape can be modelled and used to generate a new coordinate frame by which to calculate each jump. If modelled correctly, when viewed in this new coordinate frame, the parameter space (and, by extension the probability distribution for each jump) will be spherically symmetric, as required (Press et al., 2007c).

For a degenerate relationship (as shown in 5.7), we can characterise the correlation between the parameters using a simple, linear approximation, such that:

$$\mathbf{Q_g} = \tau \mathbf{Q_h} \mathbf{k} + \phi \tag{5.20}$$

where $\mathbf{Q_g}$ is the distribution of accepted trials for the parameter g; $\mathbf{Q_h}$ is the distribution of accepted trials for a different parameter, h (where $g \neq h$); τ is the gradient of a linear fit to their correlation; and ϕ is the constant term of this

relationship.

Since we wish to transform parameter space in such a way as to remove this correlation, our new coordinate system would need one axis parallel with this relationship and the other perpendicular to it. As such, to transform our current coordinate system into this prefered set of axes, we construct a matrix to rotate the parameter set from one reference frame to the other. The required angle of rotation can be simply calculated as:

$$\Theta_{g,h} = \tan^{-1}(\tau) \tag{5.21}$$

where $\Theta_{g,h}$ is the angle of rotation for the relationship between parameters g and h; and τ is the gradient of the linear fit to that relationship.

We can therefore create our rotation matrix with the form:

$$\mathbf{\Pi_{g,h}} = \begin{pmatrix} \cos\Theta_{g,h} & -\sin\Theta_{g,h} \\ \sin\Theta_{g,h} & \cos\Theta_{g,h} \end{pmatrix} \tag{5.22}$$

where $\mathbf{\Pi_{g,h}}$ is the rotation matrix transforming the parameters g and h into a coordinate frame where their relationship is spherically symmetric; and $\Theta_{g,h}$ is the previously calculated rotation angle between those two coordinate frames.

Figure 5.9 shows the same set of trials for two parameters in both the standard and rotated coordinate frames. Note that, before any rotation is performed, the

existing parameter-set must be normalised so that the origin of its current axis-set lies at the mean value of the distribution and to maintain the same standard deviation between points through the transformation.

Calculating the rotation angle for every combination of parameters used to construct our model light curve allows the creation of a composite $n \times n$ rotation matrix (where n is the number of parameters defining the model light curve, Γ). The inverse of this matrix can then be used to convert the initial spherically symmetric probability distribution into the same coordinate frame as the parameter degeneracies. Applying this matrix to each new trial will rotate the probability distribution of the step into this new coordinate basis where parameter space can be considered spherically symmetric comparative to the parameter dependencies.

$$\mathbf{Q'_k} = \mathbf{Q_k} \wedge \mathbf{\Pi^{-1}} \tag{5.23}$$

where $\mathbf{Q_k}$ is the new parameter set generated for the current (kth) iteration, as per Equation 5.4.1.3; $\mathbf{\Pi^{-1}}$ is the complete, inverted $n \times n$ rotation matrix for all n parameters; and $\mathbf{Q'_k}$ is the rotated parameter set which can then be used to generate the model light curve (Γ_k) for this iteration.

As with the parameter rescaling (see Section 5.4.1.3) it is difficult to precisely quantify the improvement attributed to this enhancement - since different data sets exhibit different levels of degeneracy. However, rotating parameter space in this manner can typically be expected to produce a $\sim 50\%$ reduction in correlation

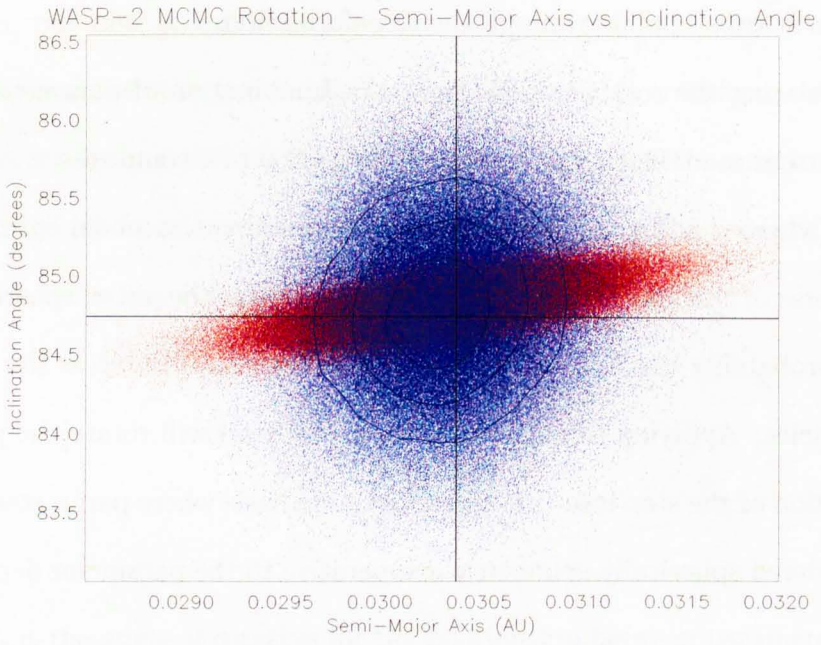


Figure 5.9: *The same distribution of accepted trials between the Inclination Angle (i) and Semi-Major Axis (a). The red points show the distribution in the standard i versus a coordinate frame (as seen in Figure 5.7), where the correlation between the parameters is distinct. The blue points show the same distribution plotted in the rotated coordinate frame where the distribution is spherically symmetric. The blue distribution has no physical meaning in the plotted parameter space and appears only for comparison. It is calculated from the red distribution by normalizing the red points to a mean of* 0 *and a standard deviation of* 1*. The normalized distribution is then rotated by application of* $\boldsymbol{\Theta}$ *and renormalized to the same centre and standard deviation as the red points. Note that no such normalization is required for the step sizes themselves (which are instead rotated by* $\boldsymbol{\Theta}^{-1}$ *from a circular distribution into the degenerate distribution) since we perform the rotation only on the step through parameter space (i.e. before it is added to the current value of the MCMC chain). Since each step is, by definition, centred on* 0 *and can be considered to have a normalized distribution as a result of the rescaling process detailed in Section 5.4.1.3, no such normalization is required.*

length.

5.4.1.5 The Final Algorithm

Combining the application of priors, step-size rescaling and parameter rotation typically reduces the correlation length of an MCMC chain to < 10 iterations - effectively improving the efficiency of the algorithm by a factor of ~ 10.

A standard MCMC run with these improvements consists of an initial burn-in period ($\sim 10^4$ iterations) to allow the chain to locate itself around the ideal solution (see Figure 5.10). To aid this process an initial modifier is applied to the test variable, γ to reduce the chance of acceptance for trials with inferior values of χ^2. This modifier is reduced every $\sim 10^3$ trials to force the chain to locate on some local minima before being allowed to begin exploring parameter space properly[1].

This step is then followed by a round of Step-Size Rescaling which continues until the desired acceptance rate of $\sim 25\%$ has been reached uniformly across all parameters. This phase typically takes $\sim 10^6$ iterations - depending on the accuracy of the initial estimates and the complexity of parameter space.

The rotation matrix, $\mathbf{\Pi^{-1}}$ is then calculated and trials continue, recalculating $\mathbf{\Pi^{-1}}$ every $\sim 10^5$ trials, until the change in the various rotation-angles (Θ) becomes negligible. This is then followed by a second round of step-size rescaling to adapt to the new coordinate frame (typically taking $\sim 10^5$ iterations). When the

[1]This allows for an independent, randomized starting point for the chain, but without the risk that such an initial condition may allow the chain to take an unfeasibly large number of iterations before locating the parameter space around the ideal solution - as could easily occur, for example, in chains with a small starting step size.

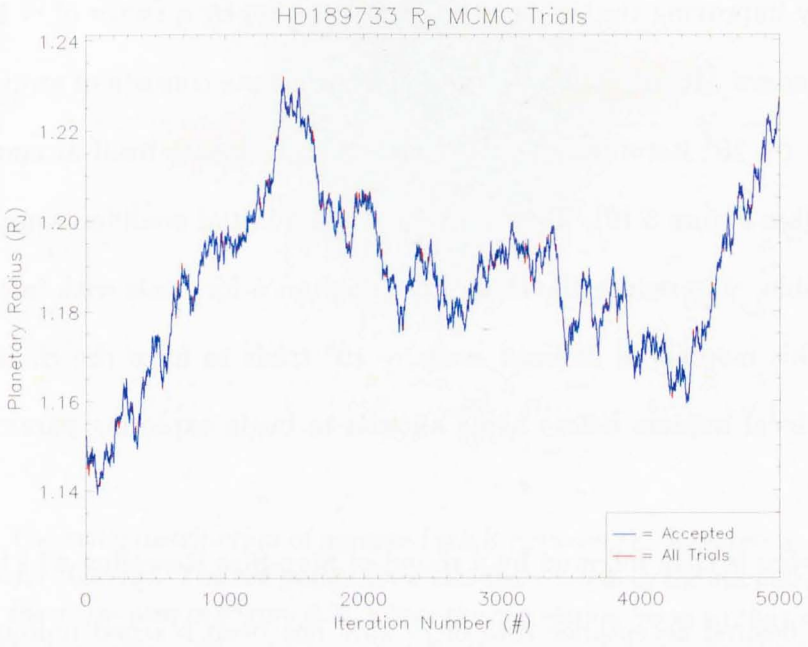


Figure 5.10: *The location of the Planetary Radius parameter (R_P) in parameter space throughout the initial burn-in period. The blue line demonstrates the accepted trials while the red line demonstrates all trials, including those which were rejected. Note that the majority of trials are currently being accepted in this chain - evidence that the step size for R_P is too small for an efficient exploration of parameter space.*

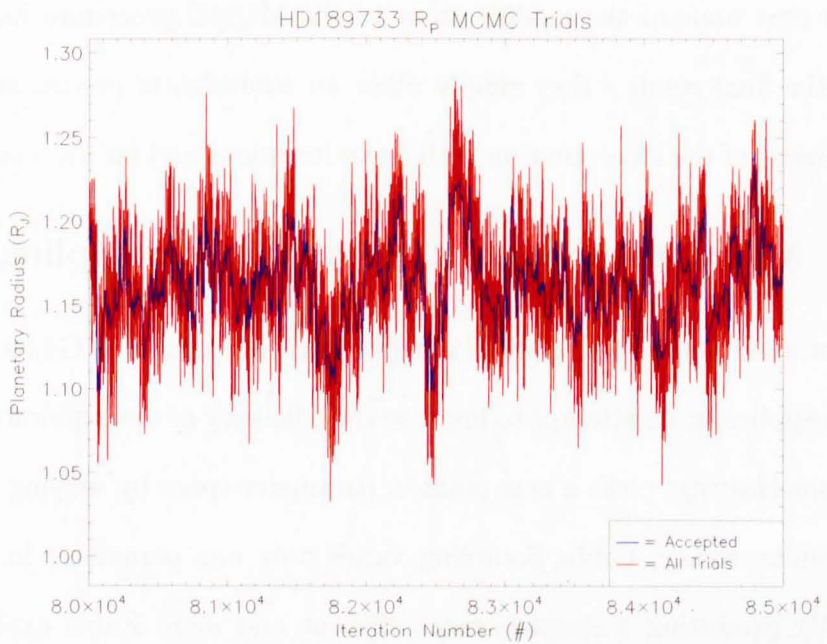


Figure 5.11: *The final stage of simulation for the Planetary Radius (R_P) chain found in Figure 5.10. Note the steady, continuous nature of the exploration - evidence that the chain is readily exploring the entirety of the available parameter space in an efficient manner.*

acceptance rate has converged once more, the parameter weightings and rotation matrix are saved (so they do not have to be recalculated for future runs) and the true MCMC procedure begins.

Only accepted trials from this last phase are used in the calculation of the ideal solution. Typically, between 10^7 and 10^8 accepted trials (see Figure 5.11) are conducted before this calculation takes place (see Section 5.5 for more detail on how this figure is determined). The standard number of iterations for the burn-in period and each recalculation, as well as the overall target of accepted trials, can all be modified via the `Config` file.

Note that none of the modifications to the MCMC procedure have any effect on the final result - they simply allow an equivalently precise and reliable measurement of the ideal solution with fewer iterations and far less computation.

5.4.2 Metropolis-Hastings versus Gibbs Sampling

An adaptation to the Metropolis-Hastings algorithm, known as Gibbs Sampling, was also applied in an attempt to improve the efficiency of the exploration. While Metropolis-Hastings picks a new point in parameter-space by varying all parameters simultaneously, Gibbs Sampling varies only one parameter in each trial - typically producing a simpler, more efficient and more stable exploration of parameter-space.

However, once the above modifications were made to the Metropolis-Hastings algorithm, the difference in efficiency between each method was found to be negligible. Both techniques were also found to produce almost identical results - though this would be expected since they are simply two different means of exploring the same parameter space.

Unfortunately, due to the single-parameter nature of Gibbs Sampling, this alternative method was found to be more sensitive to inaccurate initial conditions - in particular, a poor starting step-size for one parameter could easily slow the entire algorithm. By contrast, the adapted Metropolis-Hastings algorithm is able to cope with very loose initial conditions and is typically able to refine its exploration of parameter-space in $< 10^6$ iterations. The decision was therefore made to forgo Gibbs Sampling in favour of the modified Metropolis-Hastings algorithm.

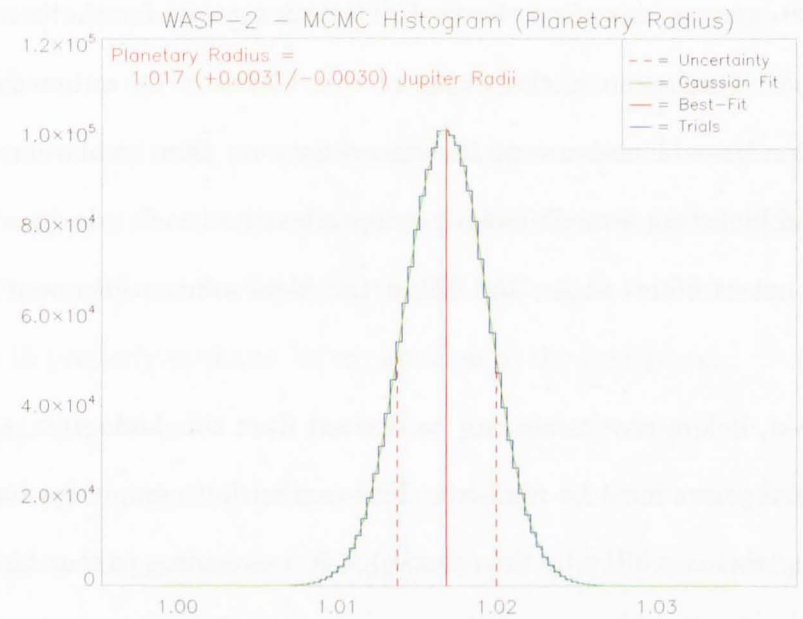


Figure 5.12: *A complete MCMC chain for a single parameter (in this case the planetary radius, R_P) binned to produce a histogram (blue) of the density of accepted trials across parameter space. The solid red line shows the derived value of the ideal solution for this parameter and the dashed red lines show the corresponding uncertainty limits on this figure. The overplotted green lines show the independently fitted semi-Gaussian curves used to measure this uncertainty in each direction through parameter-space.*

5.5 Histogram Fitting

Once a complete MCMC chain has been amassed, the parameters of the accepted trials must be processed to derive the ideal solution for the model light curve, Γ. If a sufficient number of accepted trials have been conducted, the values for each parameter can be binned into a histogram to determine the relative density of accepted trials across parameter-space. Such a histogram is shown in Figure 5.12.

Typically, such a histogram will exibit a skewed, Gaussian shape, with the

peak of the curve representing the best-fit solution, $I(j)$, for the parameter in question (j). Furthermore, the width of this Gaussian on either side of the ideal solution gives a measure of the uncertainty on that ideal value and can be measured by fitting semi-Gaussians independently to each side (producing independent uncertainties above and below the ideal solution, denoted as $\Phi^{+}(j)$ and $\Phi^{-}(j)$).

Note that, before any values can be derived from this histogram, the effects of the applied priors must be removed. This can be done simply by dividing the number of trials in each bin by the prior value corresponding to that bin's central value.

In effect, the histogram can be described as:

$$H_j(l) = \begin{cases} \left(\frac{\Psi(j)}{\Omega(E_j(l))}\right) & \text{where } E_j(l) = I(j) \\ \left(\frac{\Psi(j)}{\Omega(E_j(l))}\right) e^{-\left(\frac{E_j(l)-I(j)}{\Phi^{-}(j)}\right)^2} & \text{where } E_j(l) < I(j) \\ \left(\frac{\Psi(j)}{\Omega(E_j(l))}\right) e^{-\left(\frac{E_j(l)-I(j)}{\Phi^{+}(j)}\right)^2} & \text{where } E_j(l) > I(j) \end{cases} \tag{5.24}$$

where $H_j(l)$ is the value of the histogram for parameter, j at bin l; $E_j(l)$ is the central value of that bin; $I(j)$ is the ideal solution for that parameter; $\Psi(j)$ is the peak value of the histogram for that parameter; $\Phi^{+}(j)$ and $\Phi^{-}(j)$ are the uncertainty estimates above and below the ideal solution respectively; and $\Omega(E_j(l))$ is the prior for that parameter calculated at the central value of each bin.

Dividing out the prior for each parameter accounts for any skewness imprinted

on the distribution of accepted trials. However, it should be noted that once the MCMC chain has optimized its exploration of parameter space (through rescaling the step-size and rotating parameter space) the prior should be so broad compared to the acceptance distribution, that the effect of the prior is negligible. Without the application of priors, however, the MCMC chain could easily become too unstable to properly optimise its exploration in the first place.

Solving the above for each parameter provides the complete ideal solution, $I(j)$, required to describe the best-fit model light curve, Γ, complete with independent uncertainty estimates, $\Phi^{\pm}(j)$ both above and below each parameter in the ideal solution.

Assuming a sufficient number of accepted trials have been performed, the values of $I(j)$ and $\Phi^{\pm}(j)$ will provide a robust, reliable model light curve. However, more importantly, if $I(j)$ and $\Phi^{\pm}(j)$ have converged on the ideal solution, any further trials would do little to change the values of either. It is therefore possible to determine the required number of trials required to precisely model any light curve by recalculating the current values of $I(j)$ and $\Phi^{\pm}$ at set intervals in the MCMC chain and continuing iterations until the resulting values exibit negligible change with each new recalculation.

5.6 Prayer Beading

The MCMC process also provides us with the opportunity to acheive a measure of the residual systematic noise in a given light curve. If we consider only the residuals of the light curve, we could describe a light curve with no systematic

noise as being simply a random scatter about a constant line. In such a light curve, the order of the points in this random scatter would have little influence on the fit to the model. However, for a light curve with residual systematic noise, the ideal solution of the MCMC chain would be affected differently; depending on which parts of the light curve were most strongly impacted. Considering this in the context of a measured light curve, it is possible to quantify residual systematic noise by examining the variation in the ideal solutionacross the residuals.

If the points on the light curve were shifted along the model generated by the ideal solution - so that different sections of the observations fell at different points on the same transit curve, the parameter space explored by the MCMC chain would be distorted by any residual systematic noise, resulting in a comparably distorted histogram of accepted trials. Including this technique (called 'Prayer Beading' Gillon et al. (2007a)) in the MCMC algorithm allows for a direct assessment and quantification of any residual systematic noise.

Once the ideal solution has converged, the residuals (R) of the fit can be isolated by subtracting the model light curve (Γ) from the measured light curve (L). From the final values of the original chain, a new chain is then begun - inheriting the step-sizes and rotation of the original chain - to explore the prayer-beaded light curve. For each trial, a new, prayer-beaded light curve is created by shifting the residuals along the model of the ideal solution before that model is added back in - in this way, the transit remains the same, but the distribution of points about that curve has changed.

This secondary chain is then continued until enough trials have been per-

formed for its own ideal solution (hereafter prayer-bead solution) to converge. As with the main trial, the step-sizes for the parameters are rescaled to achieve an acceptance rate of 25% - because the parameter space explored during Prayer Beading is different to that previously explored, the step size for each parameter must be scaled once again to ensure an efficient and comprehensive exploration. However, since the transit remains the same, it is not necessary to recalculate the rotation matrix for the new chain - a spherical exploration of parameter space can still be acheived by application of the original chain's rotation matrix. As shown in Figure 5.13, the prayer bead histograms become narrower than their predecessors - revealing the impact of any residual systematic noise on the ideal solution.

This narrowing highlights the different levels of residual systematic noise present in the in- and out-of-transit points. Since our systematic model is extrapolated into the in-transit parameter space, it is inevitable that the in-transit points will be more greatly affected by residual systematics when compared to the out-of-transit points. The ideal solution of the MCMC chain is dependent only on the in-transit points (with the normalized out-of-transit model flux fixed at unity) and so can be considered an upper-limit on the residual systematic noise. During prayer-beading, the solution of MCMC chain becomes dependent on a mix of in- and out-of-transit points which can be considered to have a better overall level of detrending and therefore produce a model with smaller uncertainties.

It should be noted, of course, that the prayer-beading solution itself cannot be considered accurate as it is resampled from the original ideal solution and can only

produce a result inferior[1] to its predecessor. That said, the impact of systematic noise on the ideal solution can be quantified by measuring the reduction in the uncertainties when compared to the Prayer Bead Solution. The simplest means of acheiving such a figure is to take the ratio of the parameter space encapsulated by the uncertainties of each solution. We define this systematic measure, S as:

$$S = \frac{\Phi^+ - \Phi^-}{\Phi^+_{PB} - \Phi^-_{PB}} \tag{5.25}$$

where S is the measure of the impact of residual systematics on the Ideal Solution; Φ^+ and Φ^- are the uncertainties on the Ideal Solution; and Φ^+_{PB} and Φ^-_{PB} are the correspondind uncertainties on the Prayer Bead Solution.

Applying this method to our previous Linear Regression and Preferential Decorrelation light curves reveals the full extent of residual systematics present in each. Note that the ΔT parameters cannot be included in the Prayer Beading trials since they deal only with the relative motion of the visits to the measured phase and so would produce exactly the same Ideal and Prayer Bead solutions. As can be seen from Table 5.1, a light curve detrended by a process of linear regression consistently displays a greater amount of residual systematic noise when compared to that same curve detrended via Preferential Decorrelation. In both cases, the residual systematics are the result of extrapolating the detrending model into the in-transit parameter space. This finding would appear to suggest

[1]In a perfect light curve, the prayer-beading result would be identical to the ideal solution and so, in theory, the prayer-beading solution can technically be said to be equal or inferior to its predecessor. However, since systematic noise is inevitable in such high-precision observations and since it is impossible to completely remove all systematic noise from the light curve (see Chapter 4, it is not possible to produce such a 'perfect' light curve.

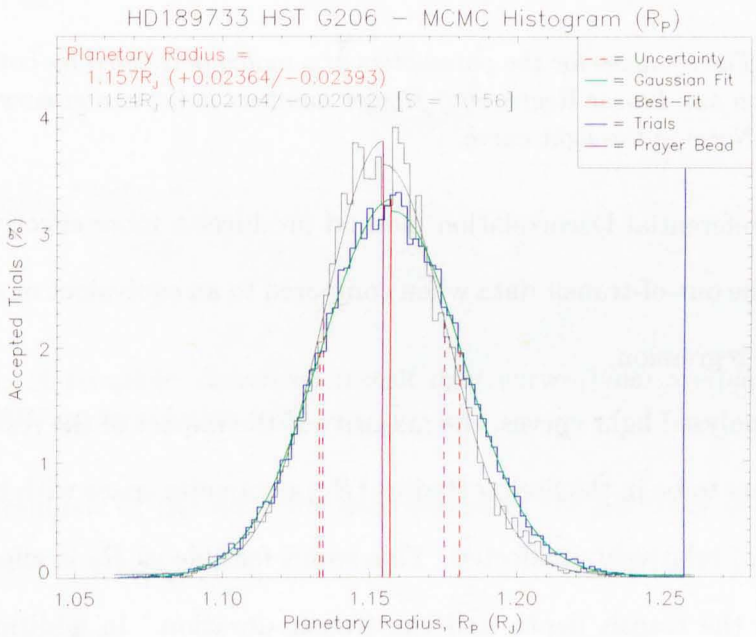


Figure 5.13: *The histogram (blue) of an Ideal Solution (red) for the Planetary Radius parameter (R_P) shown with a corresponding set of prayer-beaded trials (grey) and the resulting solution (purple).*

Model Parameter	Preferential Decorrelation	Linear Regression
R_P	1.190	1.233
R_S	1.581	1.576
i	1.323	1.356
a	1.408	1.477
U_0	1.285	1.291
U_1	1.284	1.280
Mean	1.3452	1.3689

Table 5.1: *The S values for the parameters of a model detrended by both Preferential Decorrelation and Linear Regression. Higher numbers indicate a greater prevalence of Systematic Noise in the light curve.*

that the Preferential Decorrelation method produces a more effective extrapolation from the out-of-transit data when compared to an equivalent model produced by a linear regression.

In all analysed light curves, the majority of the impact of the residual systematics appears to be in the Stellar Radius (R_S) parameter space with the Planetary Radius (R_P) relatively unaffected. This seems feasible as R_S is affected by both changes in the transit depth and the transit duration. In addition, the Limb Darkening Parameters (U_0 and U_1) typically emerge with very similar S values - the result of their impact on the light curve being of comparable scale.

5.7 The Modelled Light Curve

With a complete set of best-fit parameters, $I(j)$, and uncertainties, $\Phi^{\pm}(j)$, the best-fit light curve, Γ, can be generated and compared with the measured light curve from the data. Such a model light curve can be found in Figure 5.14, overplotted on the data set from which it was derived. Furthermore, Figure 5.15

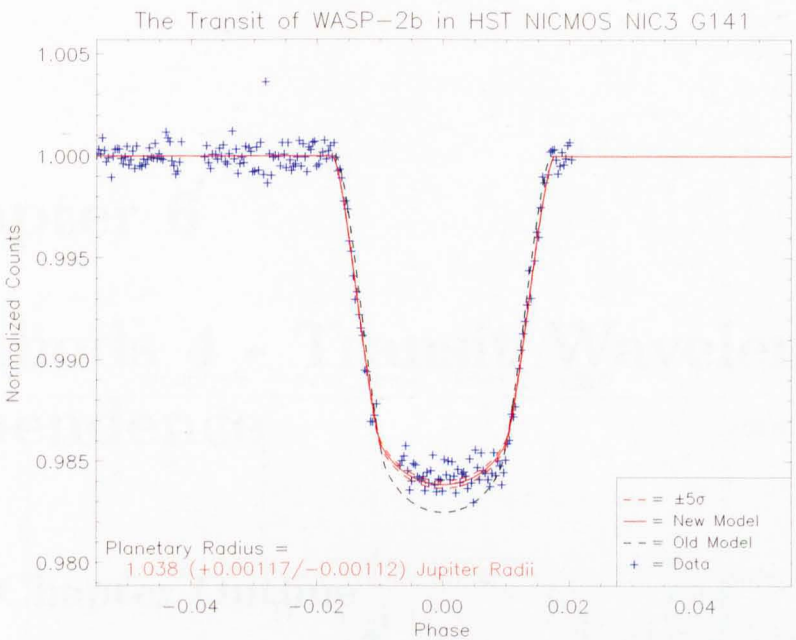


Figure 5.14: *A complete, detrended transit light curve (blue) overplotted with its best-fit, MCMC model solution (red).*

shows the residuals of the same fit, demonstrating the accuracy of the MCMC solution.

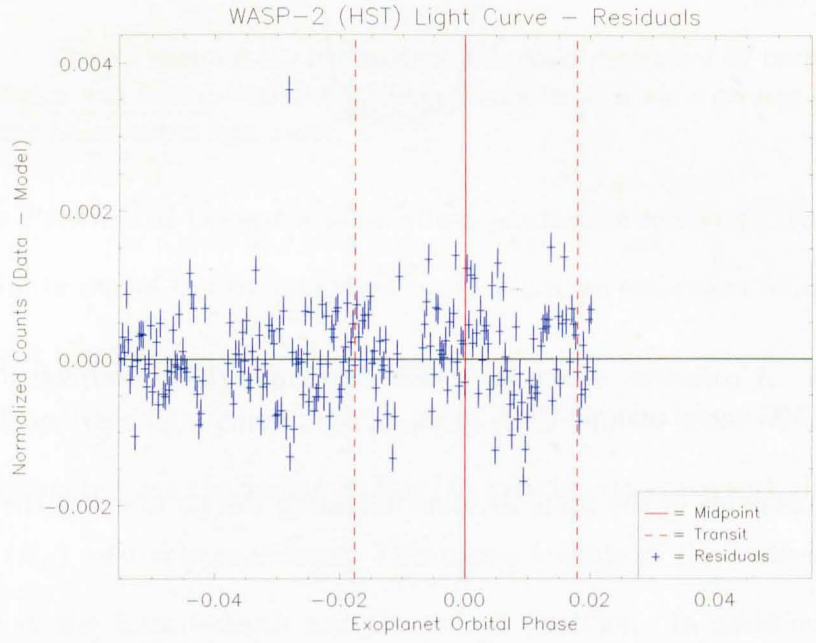


Figure 5.15: *The residuals, $(L(i) - \Gamma(i)$, blue) of the measured light curve $(L(i))$ and model light curve $(\Gamma(i))$ shown in Figure 5.14. The dashed red lines show the boundaries of the transit as calculated in Section 3.7.1.*

Chapter 6

Methods 4 - Transit Wavelength Dependence

6.1 Chapter Outline

The following chapter details the extraction and analysis of wavelength-dependent light curves from a HST NICMOS NIC3 Grism data set, in contrast to the white light curves processed thus far. Details of the various means by which such light curves can be extracted from the data are followed by the programming steps required to process these light curves prior to modeling. This is followed by an overview of the variations on the MCMC modeling technique (see Chapter 5) required to parameterise the wavelength-dependent light curves. The chapter concludes with details on the atmospheric transmission spectra derived from such wavelength-dependent analysis.

6.2 Wavelength Dependence

While the overall light curve tells us the planet's radius and orbital properties, it can tell us nothing of the atmosphere of the planet which would lend a far

greater context to those parameters. In order to determine the make-up of the atmosphere, we must assess the transit light curve at different wavelengths; specifically searching for any changes in the transit depth.

A change in transit depth would indicate a relative change in the apparent radius of the planet (R_P), implying that the planet was absorbing a different amount light at that wavelength (the effect of a change in R_P on the depth of transit can be found in Figures 5.2 and 1.13). Analysing this level of absorption across the full wavelength range of a NICMOS grism would, theoretically, produce the transmission spectrum of the planet's atmosphere (see Chapter 1 for a more detailed explanation).

6.3 Binning

The simplest means of determining the transit depth at different wavelengths is to return to our light-gathering box (see Section 3.6) and divide that box into various bins; thereby separating the first order spectrum into specific wavelength ranges. Gathering the light in these bins individually produces several wavelength-dependent light curves, like the one found in Figure 6.1.

The width of the bins used to gather light for each wavelength-dependent light curve directly governs the quality of the overall results. The smaller the bins, the more the first order spectrum can be divided and so the greater the wavelength resolution of the resulting transmission spectrum. However, as the bin size reduces, the amount of light in each bin reduces in turn; thereby increasing the RMS Spread of the resulting light curves and so increasing the uncertainties

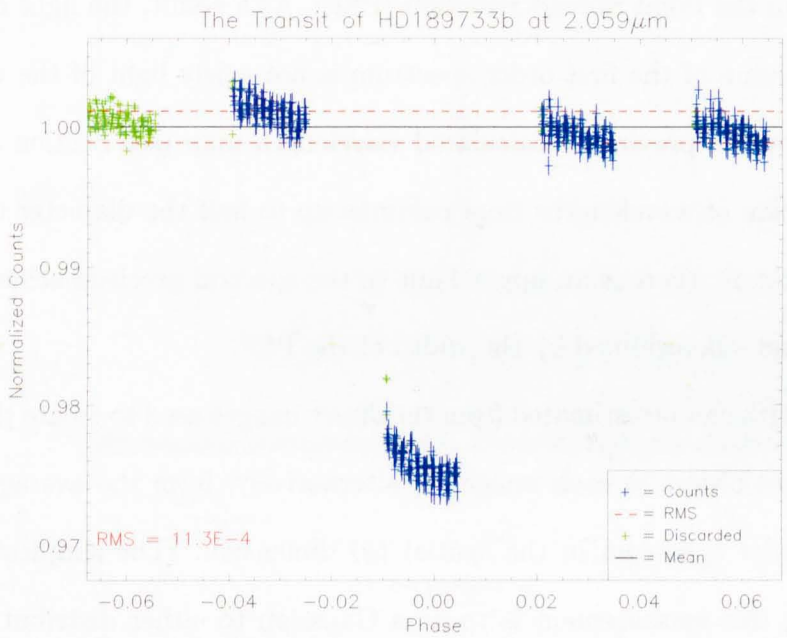


Figure 6.1: *A raw (pre-detrending) wavelength-dependent light curve (blue) binned from the HD189733 NICMOS NIC3 G141 data set using* 10% *of the columns in the first order spectrum and centered on a wavelength of* $2.059\mu m$.

on the resulting measurement of the properties of the transit light curve (in particular, of the Planetary Radius, R_P). In order to determine the optimum bin-size, we must consider its effects on both photometric and spectral precision.

However, because each NICMOS exposure is defocussed (see Chapter 2), the light from any one wavelength will be spread over several columns; depending on the width of the Point Spread Function (PSF). As a result, the light recorded in any one column of the first order spectrum is not solely light of the wavelength predicted by the previously calculated wavelength map (see Section 3.4.1), but a combination of wavelengths from columns up to half the diameter of the PSF away. Therefore, there is an upper limit to the spectral precison achievable with each data-set - determined by the width of the PSF.

This width can be estimated from the direct images used to locate the position of the target object in each image or, alternatively, from the average width of the first order spectrum in the spatial (y) dimension. The simplest means of conducting this measurement is to fit a Gaussian to either distribution and to measure the full width half maximum of the curve. Doubling this value gives an accurate estimate of the width of the PSF, as seen in Figure 6.2.

Knowing the width of the PSF provides a lower limit on the width of each wavelength-gathering bin, but does not necessarily indicate the optimum bin size to acheive the best possible trade-off with photometric resolution. To determine this optimum point, we must compare the resulting RMS Spreads of a series of wavelength-dependent light curves created from different bin sizes, as seen in Figure 6.3.

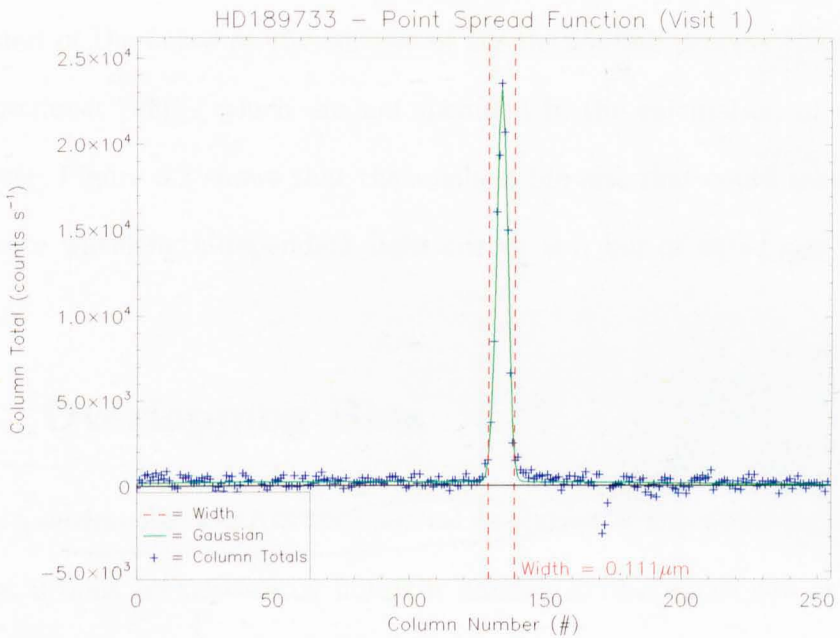


Figure 6.2: *The width of the Point Spread Function measured from the direct images taken as part of a set of HST NICMOS NIC3 images. The blue points show the column-totals of the composite direct image. The overplotted green curve shows the fitted Gaussian to the distribution and the dashed red lines show the limits of the resulting PSF width, calculated as twice the full width half maximum.*

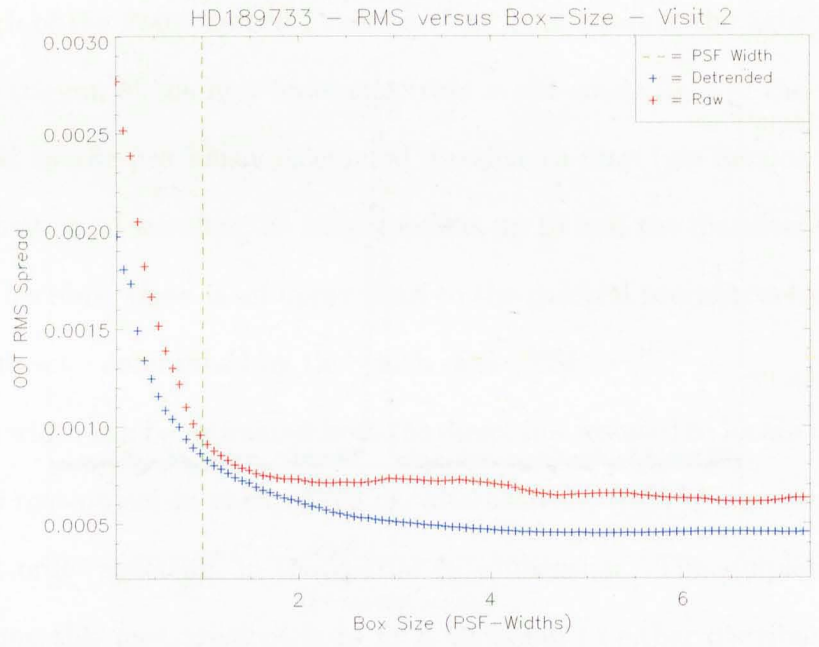


Figure 6.3: *A plot of the Normalised RMS Spreads, calculated both before (red) and after (blue) detrending, of various wavelength-dependent light curves against the size of the light-gathering bin used to create them. The dashed green line shows the lower limit on bin size determined by the width of the PSF.*

Normalizing the bin size to units of the PSF Width clearly shows a swift reduction in the quality of the light curve once the bin width drops below one PSF width (or below twice the full-width half-maximum of the PSF). For larger bins, the RMS spread of the resulting detrended light curves remains mostly constant - though, of course, the RMS spread alone does not completely represent the precision of the curve or the success of the detrending process (particularly for the in-transit points which are not included in the calculation of the spread). However, Figure 6.3 shows that the smallest bin size that could reliably be used to create wavelength-dependent light curves is a bin of equal size to the PSF itself.

6.4 Overlapping Bins

While a set bin size controls the spectral resolution of any wavelength-dependent results, it does not necessarily limit the number of bins which can be used across the first order spectrum. Overlapping a series of bins across the full length of the first order spectrum (so that the light falling in any one column may be counted in several bins concurrently) can give a better indication of the wavelength dependence of the result than simply dividing up the first order spectrum depending on the number of times our optimum bin size divides into its length.

For example, for a hypothetical transmission spectrum with a single peak, the location where the various light-gathering bins centre around that peak could serve to either accentuate or diminish its magnitude to the eye. Any bin centered directly on that peak would produce a wavelength-dependent light curve with a

far larger difference in transit depth when compared to two bins on either side of such a peak, each collecting only half of the light falling in the appropriate columns.

Furthermore, this approach can serve to make any residual systematic effects in the transmission spectrum more readily identifiable with the naked eye. Since no real feature in the final spectrum could have a width thinner than that of the PSF, any feature with such a width can only be the result of some effect of the telescope's optics or detector (i.e. something occuring after the defocussing of the incident light).

The simplest means of defining these overlapping bins is to start with the first such bin at the very edge of the first order spectrum (as defined by the overall light-gathering box, see Section 3.6) and start each new bin at one-column intervals until a bin reaches the other end of the box. Since there is no further positional information available on the sub-pixel level, there is little more to be gained by making the separation between each bin smaller than one pixel-width (i.e. one column).

Once complete, this process will typically produce over one hundred wavelength dependent light curves, each identified by the central wavelength of the bin from which they were derived (as calculated from the previously defined wavelength map, see Section 3.4.1). Since each bin remains at the optimum width, each retains an RMS spread similar to that seen in Figure 6.1.

However, at the very edges of the first order spectrum, the smaller number of photons falling in those columns can cause the RMS spread of the resulting

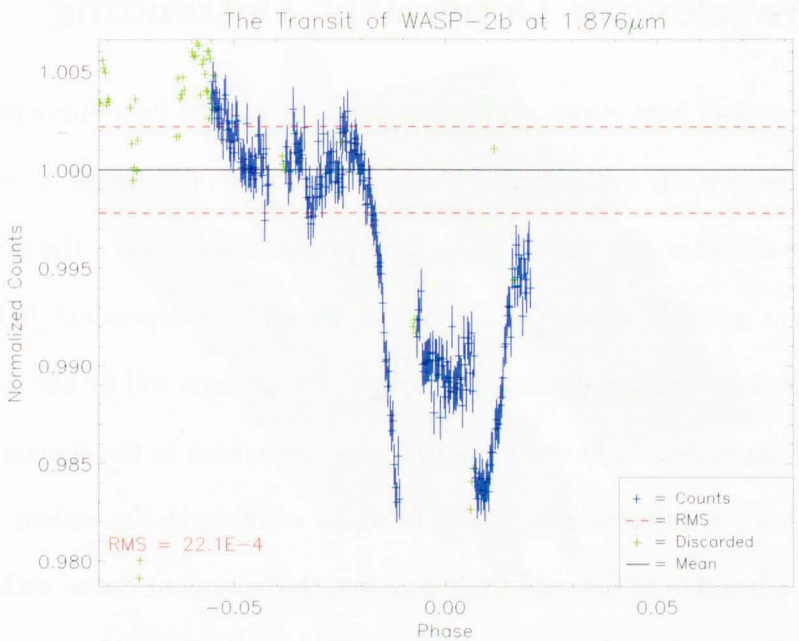


Figure 6.4: *A wavelength-dependent light curve like that seen in Figure 6.1 but binned at the very edge of the first order spectrum at 1.87μm. Note the increased spread of the points and the dominant systematic effects plaguing the data. Even once detrended, any results derived from such a light curve could not be believed.*

light curves to become undesirably high (see Figure 6.4). As such, it is best to determine a threshold of believability to avoid wasting processing time detrending and modeling light curves whose resulting parameters would be subject to undesirably large uncertainties.

Thankfully, such a threshold has been previously defined as the usable range of Gaussian fits used to calculate the detrending parameters in Chapter 4 (see Section 4.2.1). Adopting this limit and excluding all bins which fall entirely outside its range leaves a large number of believable wavelength-dependent light curves with reasonable RMS spreads.

6.5 Wavelength Dependent Detrending

As with the overall light curve, each wavelength-dependent light curve must be detrended to remove any systematic noise before it can be fitted and its parameters derived (see Chapter 4). However, the parameters which govern the overall light curve are not necessarily applicable to the wavelength-dependent light curves. For example, an overall movement in the spectral (x) axis will be the same for all wavelength-dependent light curves, however, a movement in the spatial (y) direction could have a different magnitude for each wavelength-dependent bin. As a result, it is advisable to detrend each wavelength-dependent curve with independently calculated values for the y-position (P_0), width (P_3) and background (P_6) parameters. Recalculating each of these parameters for specific bins can be easily done from the initial Gaussian parameters from which the overall parameters were calculated. As such:

$$P_0(\lambda) = \frac{\sum_{x=\mu}^{\nu}(C(x))}{\nu - \mu} \tag{6.1}$$

$$P_3(\lambda) = \frac{\sum_{x=\mu}^{\nu}(W_x)}{\nu - \mu} \tag{6.2}$$

$$P_6(\lambda) = \frac{\sum_{x=\mu}^{\nu}(a_x C(x)^2 + b_x C(x) + c_x)}{\nu - \mu} \tag{6.3}$$

where $P_0(\lambda)$ is the y-position parameter for a wavelength bin centered on λ; $P_3(\lambda)$ is the width parameter for a wavelength bin centered on λ; $P_6(\lambda)$ is the

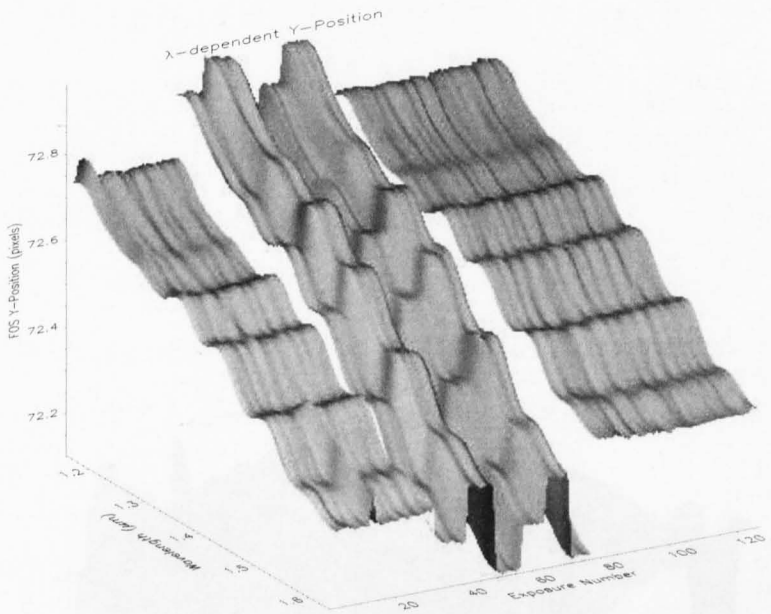


Figure 6.5: *A surface plot showing the variation of the spatial position (P_0) parameter with both wavelength (x-axis) and image number (y-axis).*

background parameter for a wavelength bin centered on λ; μ and ν are the upper and lower column numbers, x, of the boundaries of that wavelength bin; $C(x)$ is the center of a Gaussian fit to the first order spectrum along a column, x; W_x is the width of that Gaussian; and a_x, b_x and c_x are the parameters of the quatratic fit to the background light in that column.

These wavelength-dependent parameters can then be used to replace their equivalent, overall parameters when detrending; leading to more comprehensive, more robust detrended wavelength-dependent curves. Examples of each wavelength-dependent parameter's variation with wavelength can be found in Figures 6.5, 6.6 and 6.7.

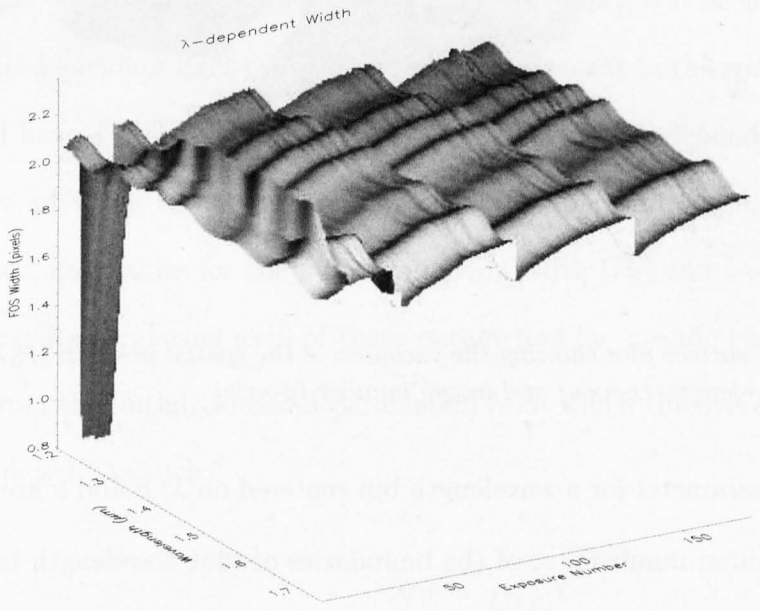


Figure 6.6: *A surface plot showing the variation of the width (P_3) parameter with both wavelength (x-axis) and image number (y-axis).*

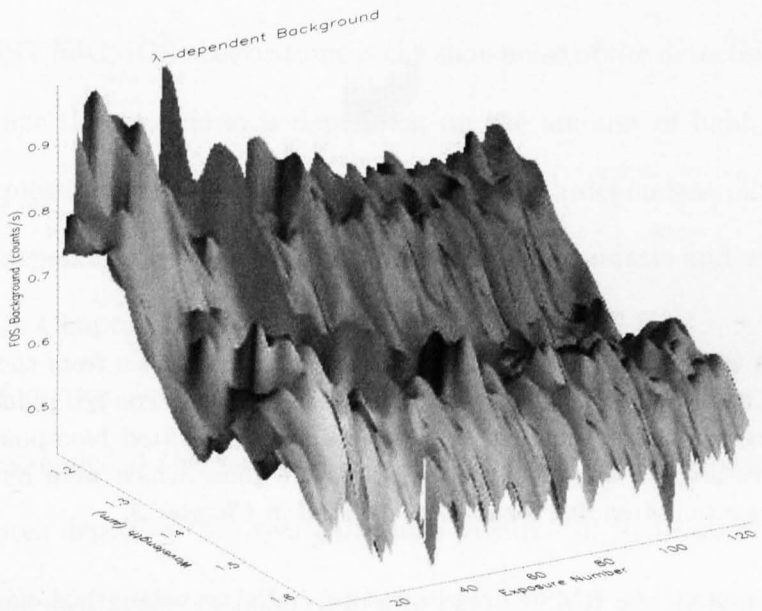


Figure 6.7: *A surface plot showing the variation of the background (P_6) parameter with both wavelength (x-axis) and image number (y-axis).*

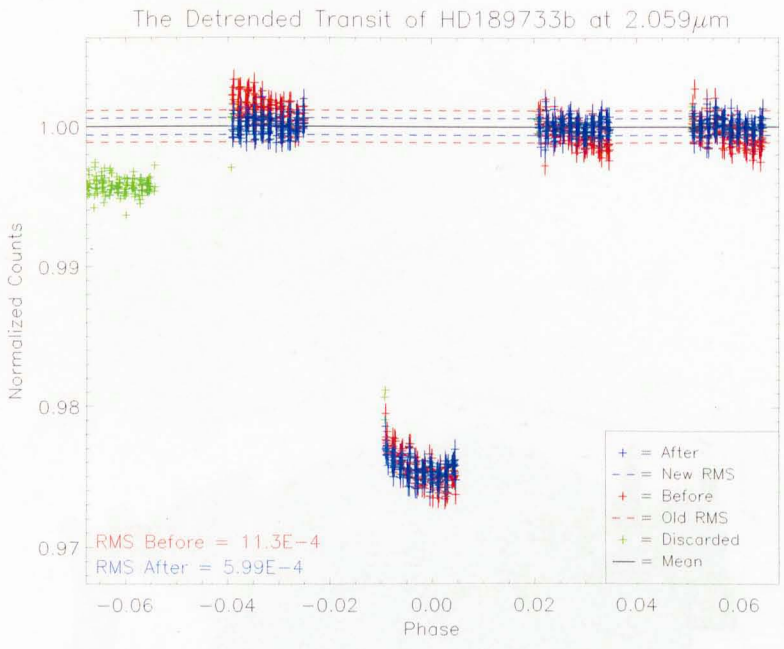


Figure 6.8: *A detrended wavelength-dependent light curve taken from the HD189733 NICMOS NIC3 G141 data set at a bin centred on 2.059μm. The red points show the light curve prior to the detrending process with the overplotted blue points showing the same curve after detrending. Green points are those which were removed from calculation prior to detrending for reasons detailed in Chapter 3.*

Once detrended, the RMS Spread of a detrended wavelength-dependent light curve typically reduces to less than the value of the non-detrended composite light curve. A typical detrended wavelength-dependent light curve is shown with reference to its previous form in Figure 6.8.

Accurate wavelength-dependent detrending is imperative, since each wavelength-dependent curve is detrended in isolation. As such, if the detrending process proves inaccurate for some proportion of the λ-dependent curves, those curves could easily retain a systematic causing variation in transit depth which would otherwise have been removed. It is therefore essential to assess the quality of

the detrending process across all wavelength-dependent curves to avoid such an occurance.

The simplest way of performing such an assessment is to analyse the RMS spreads of the detrended λ-dependent curves. If all significant systematic effects have been removed in the detrending process, the RMS Spread should be dependent only upon white (entirely random) noise. The dominant source of white noise in HST NICMOS observations is the shot-noise of the detector (Viana et al., 2009). Since the shot-noise is dependent on the amount of light falling on the pixels in question, the white noise in the wavelength-dependent curves should be similarly dependent (more discussion on residual systematic and white noise can be found in Chapter 4).

Therefore, we can consider the wavelength-dependent curves to have been reliably detrended if the variation of the RMS spread with wavelength exhibits only features dependent on real variations in flux. In these observations, there are only two sources of variation with wavelength: the stellar spectrum and the wavelength-sensitivity of the detector (since we are measuring the RMS spread from the out-of-transit light alone, we do not need to concern ourselves with any wavelength-dependence related to the transit). We can approximate both of these effects by taking the column totals of the first order spectrum across all out-of transit images. The resulting counts spectrum, $C_{OOT}(\lambda)$, can then be compared with the RMS spreads to assess any residual systematic effects.

Figure 6.9 shows the RMS spread of a series of wavelength-dependent light curves both before and after detrending. The overplotted lines represent the

(inverted) counts spectrum scaled to be shown relative to the RMS Spread[1]. The features in the undetrended RMS plot demonstrate the impact of systematic noise in each wavelength bin and, as the post-detrending points show, those features are not fully removed in the detrending process (though they are improved).

Analysis of this plot for each data set allows a more comprehensive conclusion to be drawn on the reliability of the resulting atmospheric spectrum. Not only can the overall degree of residual systematic noise be used to gauge the reliability of the results, but by asssessing the features in the RMS plots it is possible to identify corresponding features in the spectra and so identify specific systematic effects in the wavelength-dependent results.

However, the RMS is not a definitive measure of the believability of the detrending process. As noted in Chapter 4 light curves with insufficient coverage of parameter space in the out-of-transit points cannot be expected to perform a reliable removal of systematic noise. Therefore, in cases with insufficient coverage, it is best to omit the entire visit from the modeling process.

6.6 Wavelength Modeling

With a complete set of wavelength-dependent light curves, we can create independent model light curves for each by application of the MCMC algorithm detailed

[1]For a perfect light curve, we would expect the residual noise after detrending to be purely Poissonian and so would compare the residual RMS spreads to a curve of $\sqrt{C_{OOT}(\lambda)}$. However, (as noted in Chapter 4) even a perfect detrending process would be unable to acheive a final spread better than $2\sqrt{C_{OOT}(\lambda)}$ and in reality, we cannot assume the detrending to be perfect. Therefore, since it is impossible to quantify the 'imperfection' of the detrending process, we can only compare the RMS spreads to a curve of $n\sqrt{C_{OOT}(\lambda)}$, where n is a scale factor chosen to best approximate the observed post-detrending RMS spread distribution.

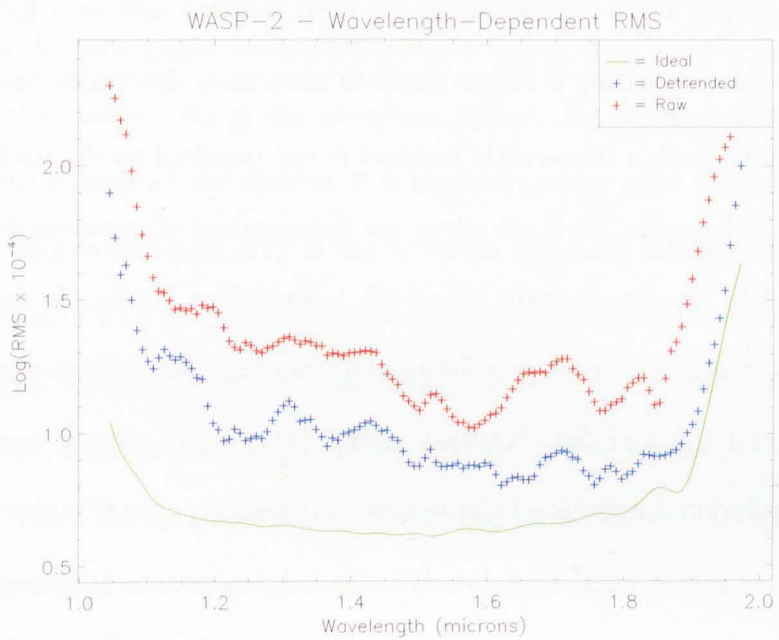


Figure 6.9: *A logarithmic plot of the RMS Spread of a set of wavelength-dependent transit light curves shown before and after a process of detrending. The undetrended points (red) show several clear features resulting from systematic effects which are universally reduced after those light curves have been detrended (blue). However, those light curves still fall short of the hypothetical 'perfect' light curve (green) calculated from the expected photon-noise at each wavelength from the detector sensitivity in this wavelength range. Analysis of such plots allows for more comprehensive conclusions to be drawn as to the volume of residual systematic noise remaining after the detrending process. This information can then help determine if features in the resulting atmospheric spectra are real, or the result of residual systemstics.*

in Chapter 5. However, processing each wavelength-dependent light curve in this manner would require a large amount of processing time - longer per curve than simply the overall light curve, as each λ-dependent curve would be expected to have a greater RMS Spread about its ideal solution and so a more complex parameter-space requiring a longer chain to accurately determine that solution.

Thankfully, such a time-costly method is not required as the majority of parameters which govern the light curve are independent of wavelength and would be expected to remain constant across all λ-dependent curves. In fact, the only parameters which require λ-dependent modeling are the limb darkening coefficients (U_0 and U_1) and the planetary radius (R_P) which should vary according to the trasmission spectrum of the planet's atmosphere (see Chapter 2).

Therefore, rather than model each λ-dependent curve in isolation, we roll all wavelength-dependent curves into the existing MCMC algorithm and model the variation of each λ-dependent parameter (R_P, U_0 and U_1) in parallel with the overall light curve.

The MCMC algorithm progresses as normal (see Chapter 5), working entirely independent of the corresponding wavelength analysis. In each iteration (after the initial burn-in period, when the overall MCMC chain has located the parameter-space containing its ideal solution), the algorithm also generates a comparative model λ-dependent light curve (Γ_λ) by inheriting the parameters from the general model (Γ) and augmenting those expected to exhibit wavelength dependence, such that:

$$\Gamma = q(R_P, R_S, i, a, U_0, U_1, \Delta T_V) \quad (6.4)$$

$$\Gamma_\lambda = q_\lambda(R_P + \Delta R_P(\lambda), R_S, i, a, U_0 + \Delta U_0(\lambda), U_1 + \Delta U_1(\lambda), \Delta T_V) \quad (6.5)$$

where, Γ is the wavelength-independent model light curve; q is the parameter set governing that model; R_P is the planetary radius; R_S is the stellar radius; i is the inclination angle of the system; a is the semi-major axis; U_0 and U_1 are the limb darkening coefficients; ΔT_V is the variation in transit midpoint for each visit of observation, V; Γ_λ is the wavelength-dependent model light curve centred on a wavelenght, λ; $\Delta R_P(\lambda)$ is the difference in transit depth between the overall light curve (Γ) and the wavelength-dependent light curve (Γ_λ) at that wavelength; and $\Delta U_0(\lambda)$ and $\Delta U_1(\lambda)$ are the corresponding variations in the limb darkening coefficients at that wavelength.

For each wavelength-dependent parameter, a perturbation is generated such that:

$$\Delta q_k(\lambda) = q_k + \xi S_P(\lambda) \quad (6.6)$$

where q_k is the value of the parameter for the overall light curve for current iteration, k; $\Delta q_k(\lambda)$ is a trial difference between that overall parameter set and its value at a wavelength, λ; ξ is a random number between -1 and 1; and $S_P(\lambda)$ is a step-size specific to each parameter in $q(\lambda)$ (i.e. ΔR_P, ΔU_0 and ΔU_1).

Generating a value of ΔP for each λ-dependent parameter at each wavelength bin creates a new λ-dependent model, Γ_λ. As with the step-size weightings in the

overall MCMC simulation (see Chapter 5), the wavelength-dependent step-sizes are also scaled during processing to arrive at an ideal value by which to acheive the most efficient exploration of the wavelength-dependent parameter space.

Each wavelength dependent model is then tested against the overall model in comparison with the corresponding λ-dependent light curve. As with the overall MCMC procedure, the test is governed by the χ^2 goodness-of-fit statistic for each model light curve (Γ and Γ_λ) in comparison with the corresponding wavelength-dependent light curve (L_λ), such that:

$$\chi^2 = \sum_{l=0}^{n-1} \left(\frac{(L_\lambda(l) - \Gamma(l))^2}{\sigma_\lambda(l)} \right) \tag{6.7}$$

$$\chi^2_\lambda = \sum_{l=0}^{n-1} \left(\frac{(L_\lambda(l) - \Gamma_\lambda(l))^2}{\sigma_\lambda(l)} \right) \tag{6.8}$$

where χ^2 is the goodness-of-fit statistic for the overall model light curve (Γ) compared to the wavelength-dependent light curve (L_λ) for a wavelength λ; χ^2_λ is the goodness-of-fit statistic for the wavelength-dependent model light curve (Γ_λ) compared to L_λ for the same wavelength; n is the number of points in the light curve; l is the index of any given point in that curve; and σ_λ is the corresponding uncertainty to each point in L_λ.

The acceptance test can then be performed as normal, such that:

$$\gamma = \frac{\overline{\Omega}_i}{\overline{\Omega}_{i-1}} \frac{e^{-\chi^2_\lambda/2}}{e^{-\chi^2/2}} \tag{6.9}$$

where γ is the test variable (see Chapter 5) or probability of acceptance; χ^2 is the goodness-of-fit statistic for the overall model light curve (Γ) against the binned light curve at a wavelength, λ; χ^2_λ is the goodness-of-fit statistic for the wavelength-dependent model light curve (Γ_λ) against that same light curve; and $\overline{\Omega}_i$ and $\overline{\Omega}_{i-1}$ are the priors (see Chapter 5) for the current and previous iterations in the overall MCMC chain.

As with the overall MCMC chain, the test variable is compared with a random number chosen from a uniform distribution between 0 and 1. If the value of γ exceeds this random number, Γ_λ is accepted for that wavelength and the value of ΔP stored for the accepted trial. In subsequent iterations, newly generated wavelength-dependent models are then tested against the last accepted Γ_λ, forming an independent MCMC chain for each wavelength bin.

Note that the prior values applied in Equation 6.6 are inherited from the overall MCMC chain and not recalculated for each wavelength-dependent curve. This is because the purpose of the priors are to prevent the MCMC chain from exploring areas of parameter space which are mathematically sound, yet physically impossible. For the purposes of the wavelength-dependent parameters, it is the overall parameter they augment which we need to constrain. Therefore, since those parameters are already constrained by the overall MCMC chain, there is no need to apply priors to the wavelength-dependent parameters.

Augmenting the overall MCMC chain in this manner ensures that the wavelength-dependent parameters inherit the uncertainties of their equivalent parameters in the overall model light curve (Γ). If the wavelength simulation was performed

separately - basing the values of $\Delta P(\lambda)$ on the pre-calculated ideal solution - the resulting uncertainties would only represent the difference between the overall and λ-dependent light curves; overestimating the true precision of the resulting spectrum.

6.7 Atmospheric Transmission Spectra

Once a sufficient number of iterations have been performed for the models of every λ-dependent light curve, the ideal solution for each can be determined by fitting a Gaussian to a histogram of the accepted trials - as is done for the ideal solution of the overall light curve (see Chapter 5). Once each ideal solution has been calculated, it is possible to examine the transit as a continuum across all observed wavelengths, as demonstrated in Figure 6.10.

In this wavelength-dependent ideal solution, the values of $\Delta R_P(\lambda)$ are analogous to the transmission spectrum of the observed planet, more commonly expressed as a ratio of the planetary and stellar radii, such that:

$$D(\lambda) = \left(\frac{\Delta R_P(\lambda) + R_P}{R_S}\right)^2 \tag{6.10}$$

where $D(\lambda)$ is the effective transit-depth[1] at a wavelength, λ; $\Delta R_P(\lambda)$ is the difference in the modelled radius of the transiting planet at that wavelength; R_P

[1]The value of $D(\lambda)$ does not necessarily equate to the actual depth of the observed transit since that value is also dependent on the inclination angle, i, and the limb darkening coefficients, U_0 and U_1. However, since we are only concerned with the wavelength dependence of the planetary radius, R_P, such effect can be ignored.

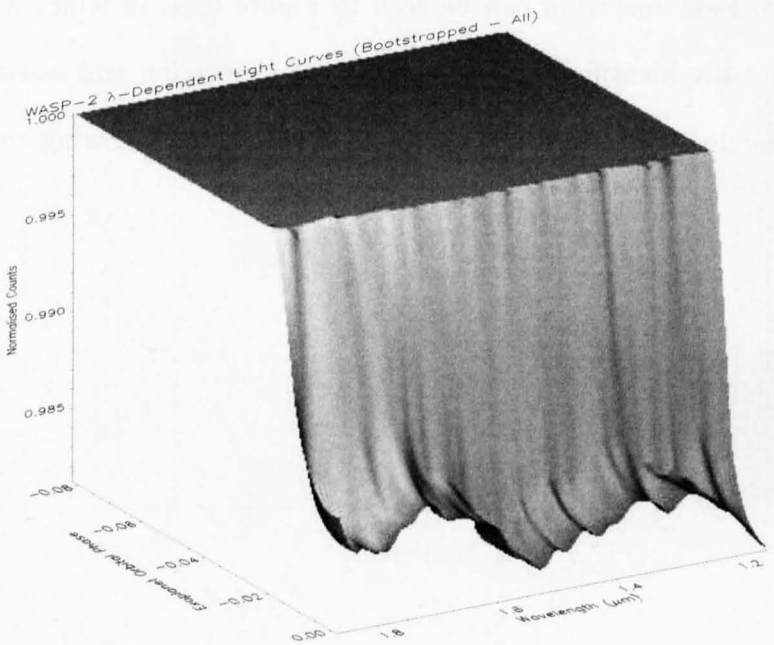


Figure 6.10: *A continuum of wavelength-dependent model light curves across a range of wavelengths, demonstrating a clear variation in transit depth.*

is the overall radius of the planet for the general ideal solution; and R_S is the radius of the parents star in that same ideal solution.

Plotting the transit depth against wavelength depicts the transmission spectrum of the planet's atmosphere - assuming the data is of sufficient quality and that a significant proportion of star light passes through the planet's atmosphere. Such a transmission spectrum can be seen in Figure 6.11, in which absorption features are readily identifiable. The specifics of identifying and assessing such features will be dealt with individually, by data set, in the following chapters.

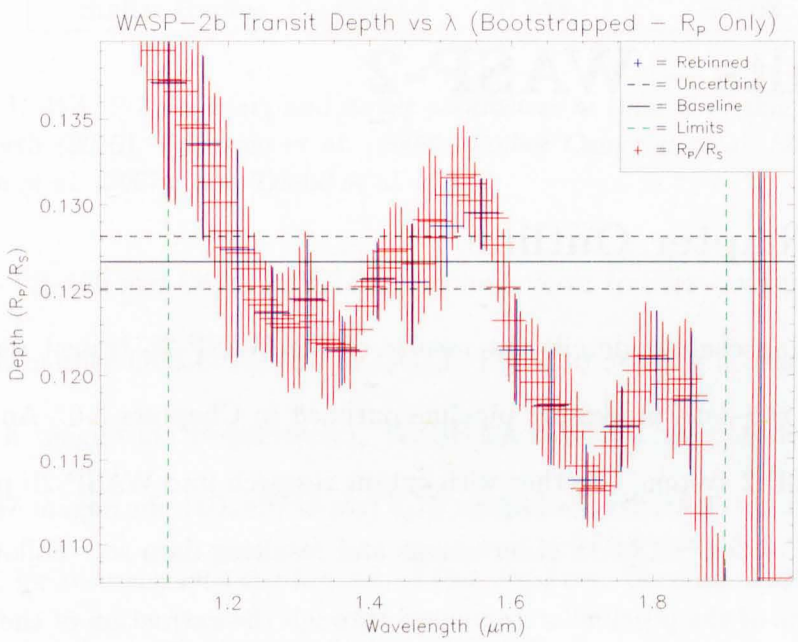


Figure 6.11: *The transmission spectrum of the atmosphere of an extrasolar planet, as revealed by a process of MCMC light curve modeling conducted on wavelength-dependent light curves extracted from a set of HST NICMOS grism observations. The red points show the transit depth for each overlapping wavelength bin, the blue points show those values rebinned to remove that overlap. The dashed green lines on either side of the plot show the believable limits of the spectrum - calculated as the width of one PSF (Point Spread Function) from either end of the first-order spectrum.*

Chapter 7

Results - WASP-2

7.1 Chapter Outline

The following chapter details the results of the WASP-2b transit data set as derived by the data processing pipeline outlined in Chapters 3-6. An overview of the WASP-2 system, together with extant research into WASP-2b proceeds a description of the NICMOS observations and resulting data set. Following this, the products of the pipeline are reviewed through the extraction of the raw light curve; the identification and removal of systematic noise; the MCMC modeling of the processed light curve; and the processing of the wavelength dependence of the transit.

7.2 Introduction

The first confirmed planetary detections from the SuperWASP project (see Chapter 1) came in 2006 with the announcement of WASP-1b and WASP-2b (Collier Cameron et al., 2007a). At the time, only twelve other transiting planets had been discovered and WASP-2b was of particular interest as only two of those planets

Parameter	Value	Uncertainty
Planetary Radius, R_P (R_J)	1.079	± 0.033
Planetary Mass, M_P (M_J)	0.847	± 0.045
Inclination Angle, i (°)	84.73	± 0.19
Semi-Major Axis, a (AU)	0.03138	± 0.011
Period, P $(days)$	2.1522254	$\pm 1.5 \times 10^{-6}$
Transit Epoch, E (HJD)	2453991.51428	± 0.0002
Stellar Radius, R_S $(R_\odot)$	0.843	± 0.08

Table 7.1: *WASP-2 planetary and stellar parameters at time of writing, sourced from Southworth (2010), Daemgen et al. (2009), Collier Cameron et al. (2007b), Collier Cameron et al. (2007a) and Triaud et al. (2010).*

(OGLE-TR-56b and OGLE-TR-113b) orbited closer to their parent stars.

With neither OGLE-TR-56b or OGLE-TR-113b orbiting a bright star (16th and 14th magnitude respectively), WASP-2's sub-12th magnitude host meant only 7.67 magnitude HD189733 and 7.75 magnitude HD209458 could rival it as a target for research into the hottest of Hot Jupiters. Because of its low density (relative to the population at the time), WASP-2b was a prime candidate for testing various theories on the nature of such bloated exoplanets. The most up-to-date parameters for the WASP-2 system can be found in Table 7.1.

Despite its extreme beginnings, WASP-2b can now be considered representative of the Hot Jupiter population as a whole. As demonstrated in Figure 7.1, WASP-2b's mass and radius lie close to the average of the population and, while a temperature of over $1300K$ (see Chapter 1 for calculation) places it far hotter than any planet in our own solar system, this is a typical figure for any Hot Jupiter.

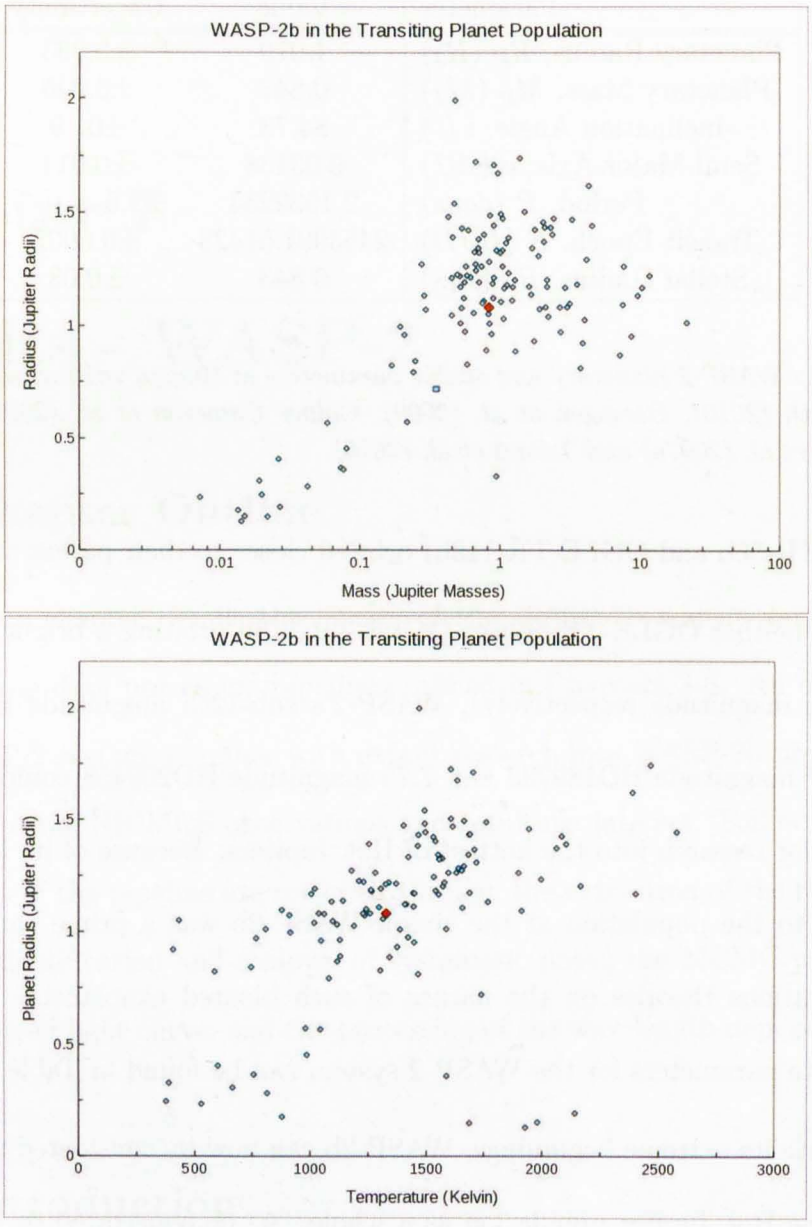


Figure 7.1: *The exoplanet WASP-2b (red) in contrast to the population of transiting planets as a whole (blue). The upper plot shows the relation between mass and radius of the planets while the lower plot shows the effect of temperature on the planets' radii. Note that in both cases, WASP-2b lies towards the centre of the distribution.*

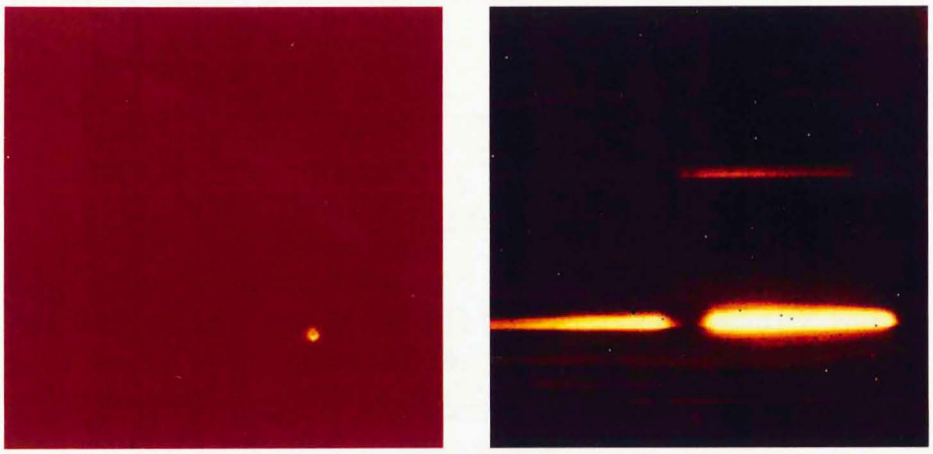
	Visit 3	Visit 4
Containing	Ingress & Egress	Transit Floor
Date of Observations	June 2007	May 2007
Number of Orbits	4	3
Filter	F166N	F166N
Filter Exposure Time (s)	21.93038	13.96211
Number of Filter Images	10	15
Grism	G141	G141
Grism Exposure Time (s)	47.99612	47.99612
Number of Grism Images	174	124
Star X Pos ($pixels$)	194.95	191.08
Star Y Pos ($pixels$)	79.40	79.71
Box X Dimensions ($pixels$)	119.12 : 243.98	115.24 : 240.10
Box Y Dimensions ($pixels$)	61.82 : 82.95	61.07 : 84.29
PSF Width ($pixels/\mu m$)	9.91 / 0.079	10.69 / 0.085

Table 7.2: *Details of the HST NICMOS NIC3 observations of the transit of WASP-2b. The upper three sections of the table hold details of the exposures and observing strategy, the lower section holds details calculated from the resulting images (as detailed in Chapter 3).*

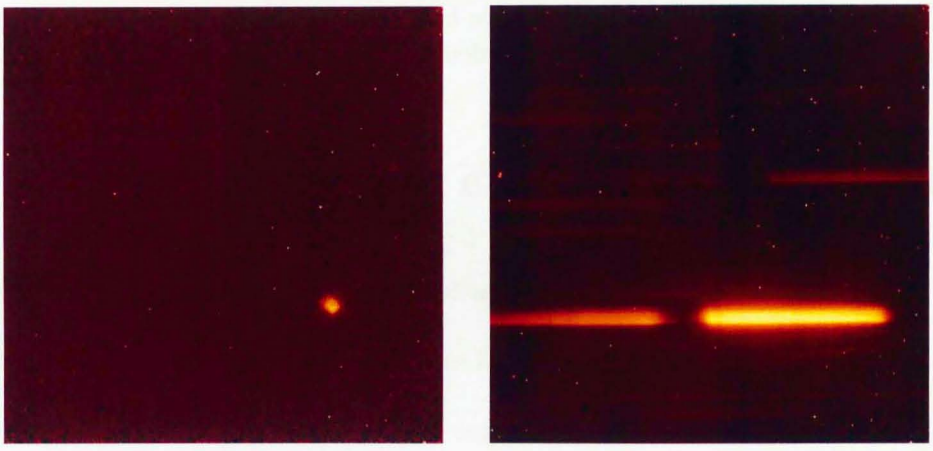
7.3 The WASP-2 Data Set

The WASP-2 HST NICMOS NIC3 data set comprises two visits of observation. In this instance, the visits are numbered 3 and 4 as they were conducted as the third and fourth visits of an HST programme which included the transit of WASP-1. Full details of each visit of observation can be found in Table 7.2. Sample direct and grism images from both visits can be found in Figure 7.2.

Analysing the direct images allowed for a precise measurement of the location of the star in the images and, from this, a wavelength-map was generated to determine the mean wavelength of light falling in each column of the first order spectrum. The location of the star in both visits is demonstrated in Figure 7.3.

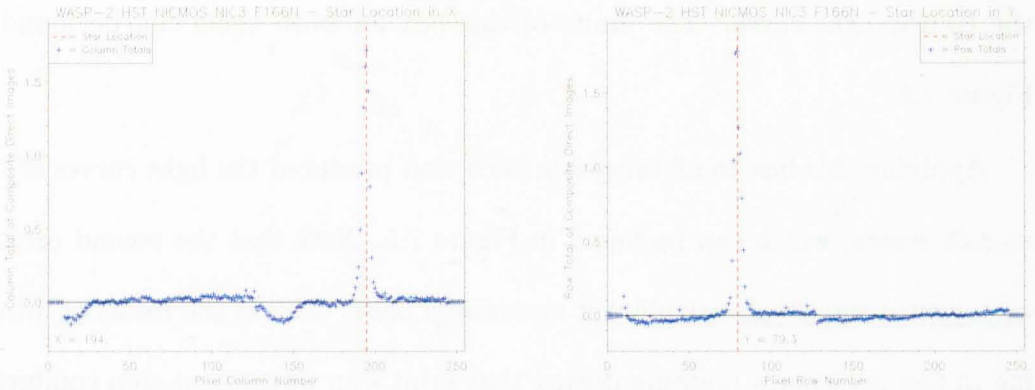


(a) Visit 3

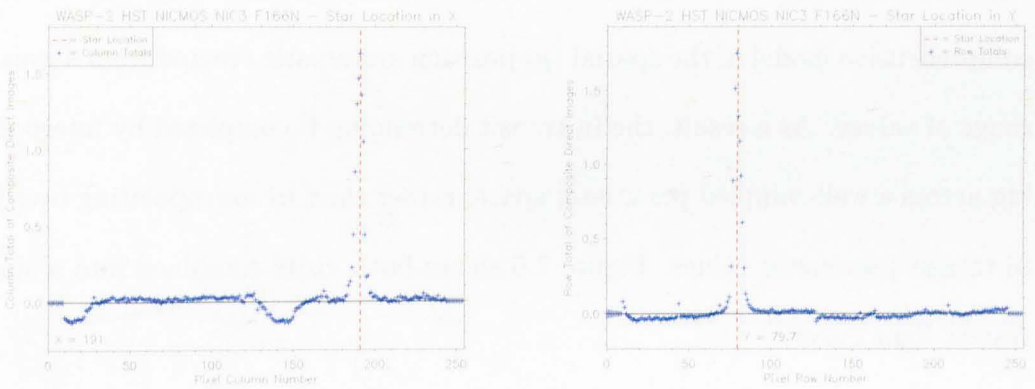


(b) Visit 4

Figure 7.2: *Sample direct (F166N) and grism (G141) images from visits 3 (top) and 4 (bottom) of the WASP-2 HST NICMOS NIC3 data set.*



(a) Visit 3



(b) Visit 4

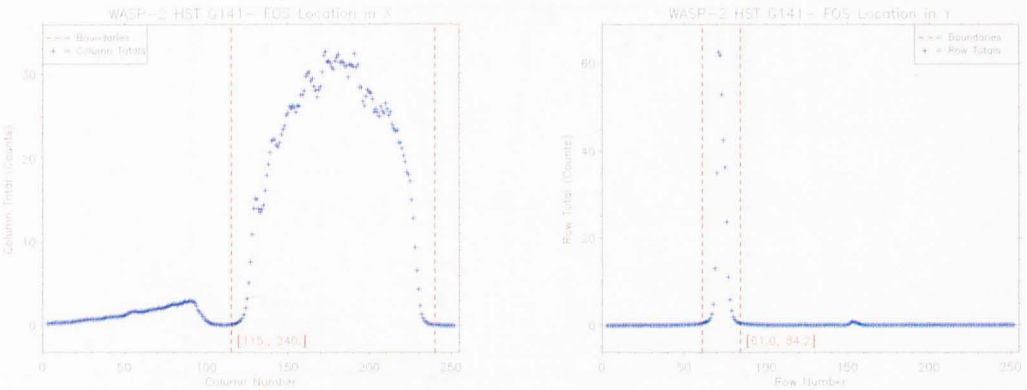
Figure 7.3: *Column- (right) and row-totals (left, blue) of a composite image comprising all exposures from each visit taken through a filter. In each plot, the location of the target star in that axis has been identified (as detailed in Chapter 3) and marked with a red dashed line. The artefacts in the background counts are features from the bias applied to the NIC3 detector - however, as only totals near to the star itself are used in the calculation of its location, they have no impact on the resulting values.*

Next, by analysing the grism images (as detailed in Chapter 3) a light-gathering box was defined around the first order spectrum by which to extract the transit light curve. The limits of this box for both visits can be found in Figure 7.4.

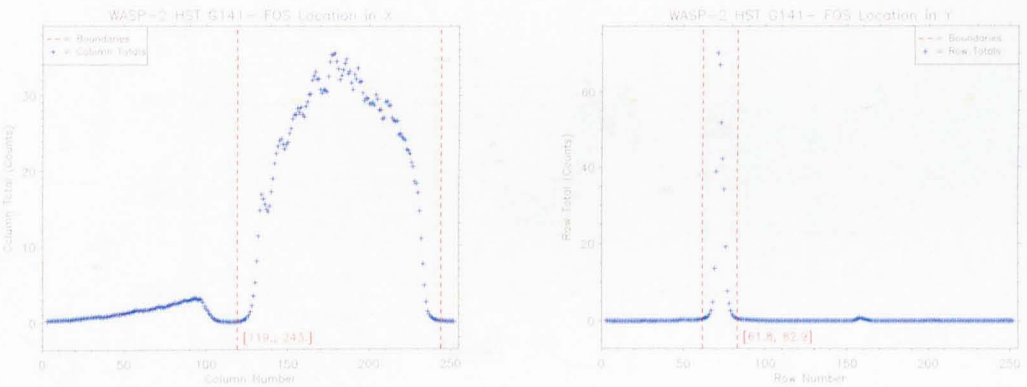
Applying this box to all images in each visit produced the light curves of two transit events, which can be found in Figure 7.5. Note that the second orbit of each visit demonstrates significant systematic noise. This is the result of dithering in the telescope's pointing during that orbit - an intentional step conducted to increase the systematic noise in the Out-Of-Transit points. This allows a more comprehensive model of the spatial (y) position systematic created from a greater range of values. As a result, the in-transit detrending is completed by interpolating across a well-sampled parameter space, rather than by extrapolating over the in-transit parameter values. Figure 7.6 shows both visits combined into a single transit light curve.

7.4 Detrending

As detailed in Chapter 4 the systematic effects present in the light curve must be modelled before they can be removed. To this end, we measure the position of the first order spectrum in each exposure of each visit, along with the temperature and temporal effects. The systematic parameters for each visit, together with their modelled impact on the recorded flux, is demonstrated in Figures 7.7 and 7.8.



(a) Visit 3



(b) Visit 4

Figure 7.4: *Column- (right) and row-totals (left, blue) of a composite image comprisng all grism exposures from each visit of observation. From these values, the dimensions of the light-gathering box (red, dashed) have been defined as described in Chapter 3. In the column plots, the main feature is the first order spectrum and the second feature (to the left of the box) is the second order spectrum. The small feature in the row plots, to the right of the spectrum, is a background star and is ignored in any calculations.*

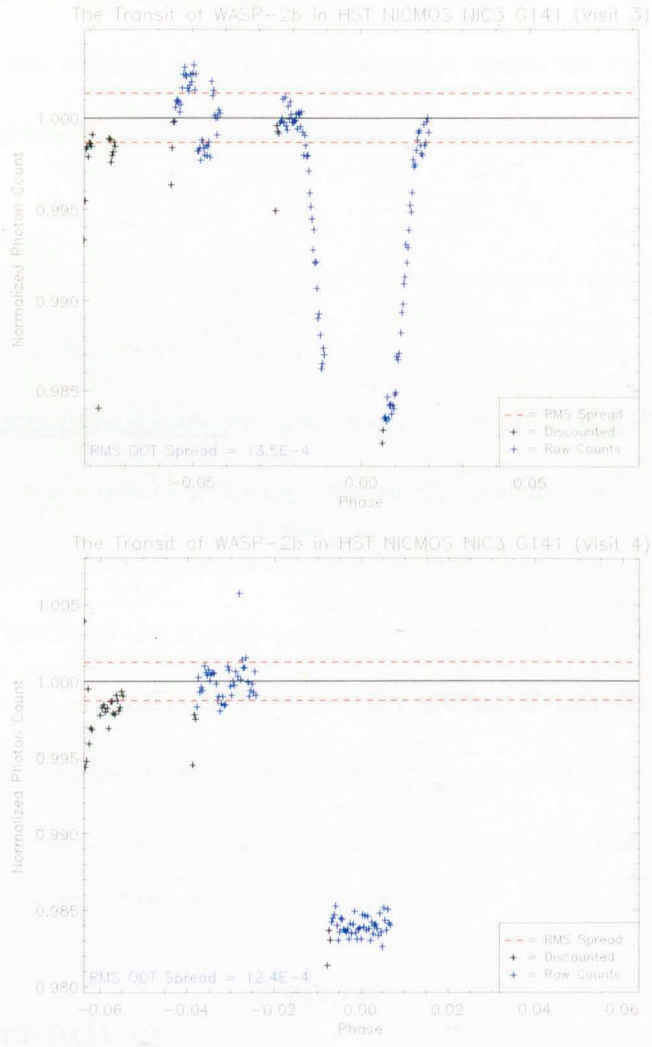


Figure 7.5: *The normalized transit light curves (blue) gathered from the first order spectrum in every exposure in visit 3 (top) and visit 4 (bottom) of the WASP-2 data set. Those points in black are from those exposures removed from processing due to expected anomalous results (see Chapter 3 for a complete explanation). Systematic effects are clearly visible in both light curves - accentuated by the dither positions visible in the second orbit in each (the first orbit in blue, since the first orbit in each visit is omitted). The red dashed lines demonstrate the range of the RMS Spread measured from the out-of-transit points.*

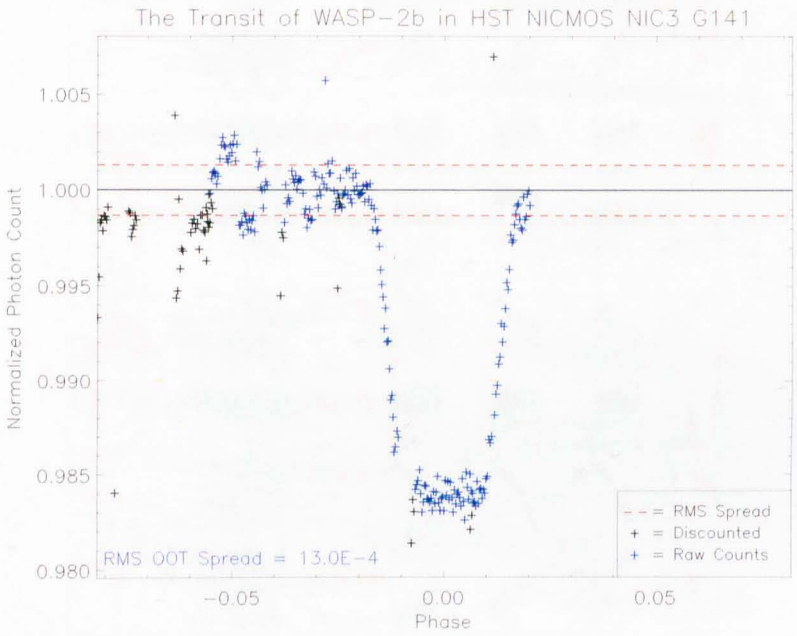


Figure 7.6: *The raw, composite transit light curve for the WASP-2 HST NICMOS NIC3 G141 data set, comprising images from both visits of observation. The black points are those images removed from calculation for various reasons detailed in Chapter 3. The overplotted dashed red lines represent the RMS Spread of the Out-Of-Transit points.*

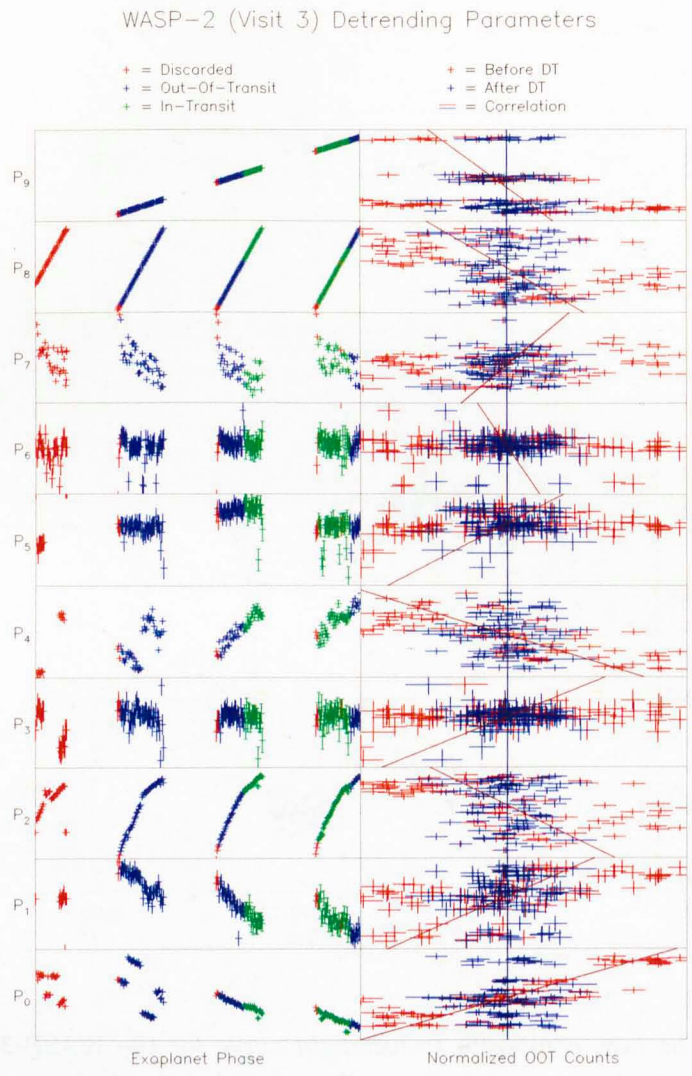


Figure 7.7: *Detrending parameters calculated for visit 3 of the WASP-2 data set. The left-hand plots show the parameters as a function of orbital phase. The red points represent exposures removed from calculation, the blue points are exposures taken out-of-transit (which will be used to model the effect of each parameter, see Chapter 4) and the green points those taken in-transit. The right-hand plots are the corresponding correlations between the parameter and the out-of-transit flux shown before (red) and after (blue) detrending. Note the distinct trends in the red points (demonstrated by the overplotted line showing the modelled relationship between the two) compared to the uniform, vertical trend in the blue.*

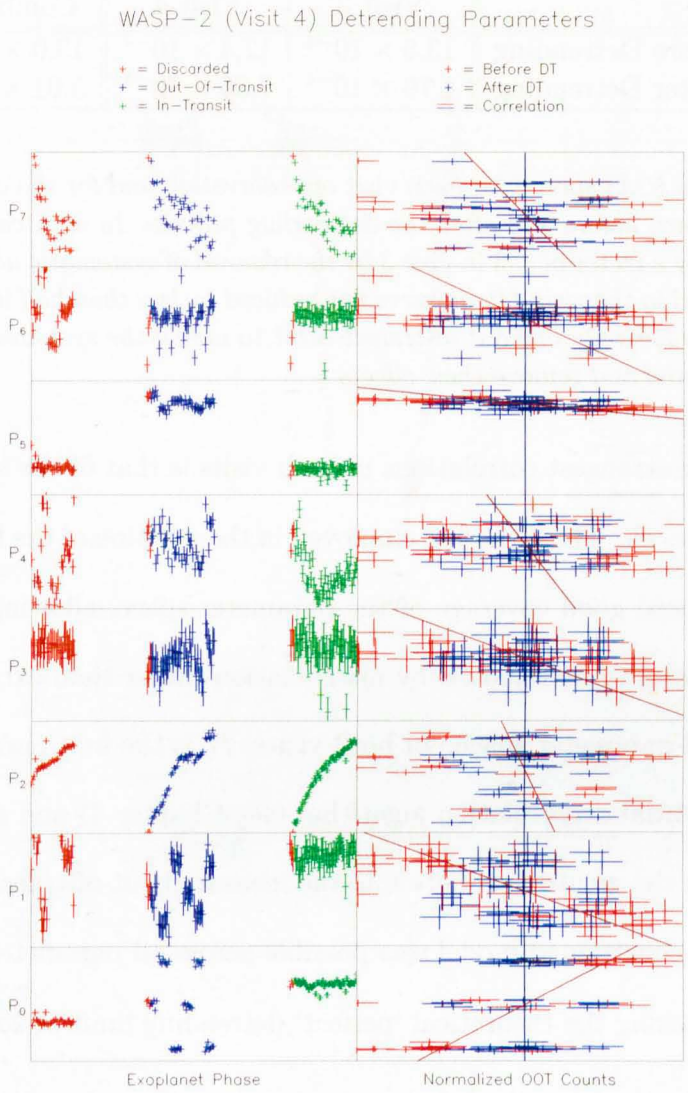


Figure 7.8: *Detrending parameters for visit 4 of the WASP-2 data set with the same layout and colour-coding as found in Figure 7.7. Note that, in this instance, the two temporal parameters (P_8 and P_9) have been omitted. This step is taken since, with only one out-of-transit orbit against which to compare, these parameters become degenerate and may result in a detrimental difference in the transit-depth. Note, in particular, the rotation parameter (P_5) in which the in-transit values are clearly not covered by the out-of-transit parameter space. As a result, it is impossible to determine whether the in-transit points have been properly detrended for this visit - and indeed, it is likely, with such a large change in rotation unaccounted for, that they have not.*

	Visit 3	Visit 4	Combined
Before Detrending	13.5×10^{-4}	12.4×10^{-4}	13.0×10^{-4}
After Detrending	4.76×10^{-4}	7.36×10^{-4}	5.01×10^{-4}

Table 7.3: *The RMS spread for each visit of observation and for the combined transit light curve shown before and after the detrending process. In each case the RMS has been reduced by a factor of* ~ 3 *in visit 3 by the removal of systematic noise. In contrast, the RMS spread in the visit 4 light curve has reduced by less than half its starting value. Since visit four has only one out-of-transit orbit to model the systematics, it is unable to properly model and remove their effects.*

Among the strongest correlations in both visits is that of the spatial (y) position parameter, P_0. The dithering employed in the position of the first order spectrum has ensured good coverage of the parameter space, allowing the in-transit systematic effects to be removed by interpolation rather than extrapolation over unconstrained parameter space. In both visits, P_0 is the first parameter selected by the preferential decorrelation algorithm (see Chapter 4) and its removal corresponds with the greatest reduction in the measured out-of-transit RMS spread. If such a well-constrained model was possible across all parameters, we could be hopeful of reaching the theoretical 'perfect' detrending limit of $2\sigma_{Poisson}$ detailed in Chapter 4.

Removing the systematic noise by a process of preferential decorrelation (as described in Chapter 4) results in a significant improvement in the RMS Spread of the out-of-transit points across both visits (exact values provided in Table 7.3). The resulting detrended light curves, shown in comparison with their raw forms, can be found in Figure 7.9.

Unfortunately, as Visit 4 contains only one OOT orbit, the detrending pa-

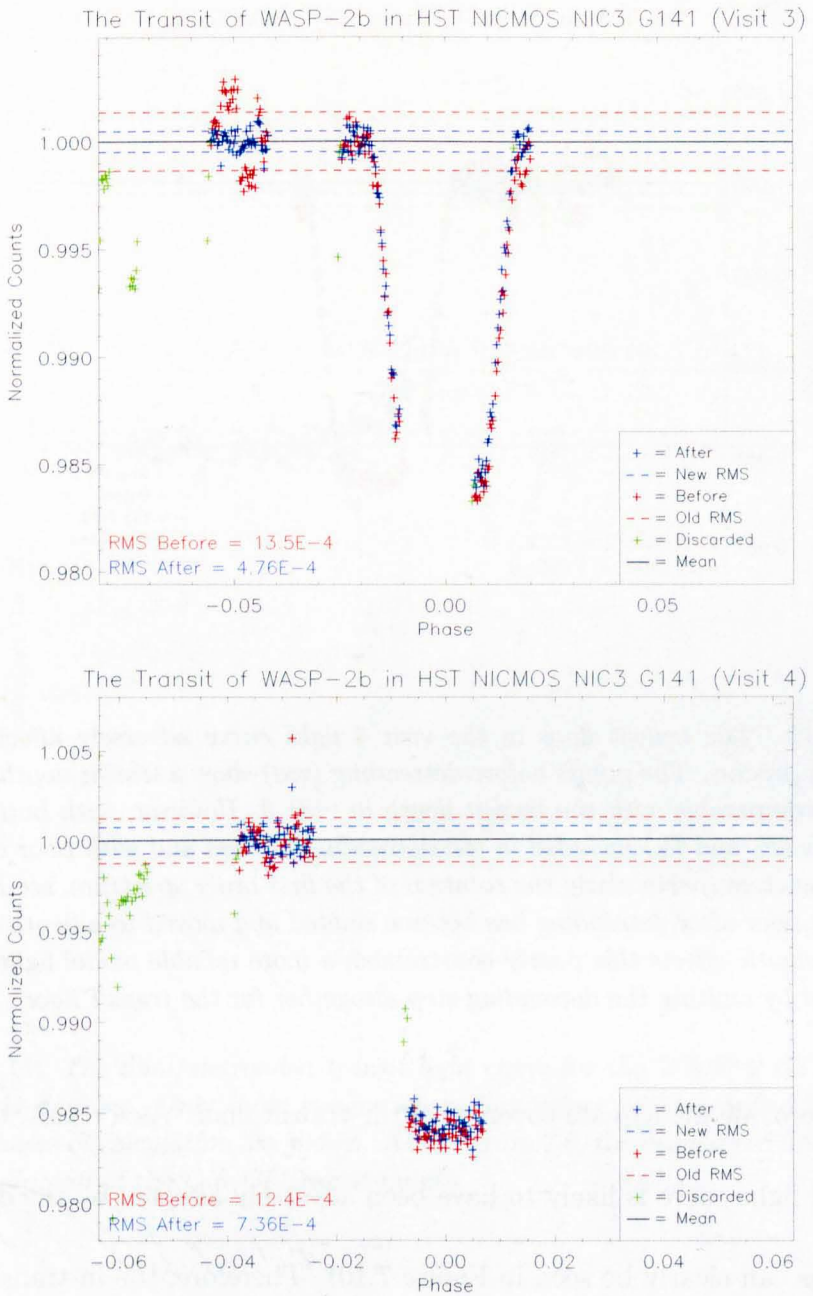


Figure 7.9: *The transit light curves from visits 3 (top) and 4 (bottom) of the WASP-2 data set shown before (red) and after (blue) detrending. The overplotted dashed lines show the relative out-of-transit RMS Spread of each light curve. Green points are those removed from processing prior to the detrending step.*

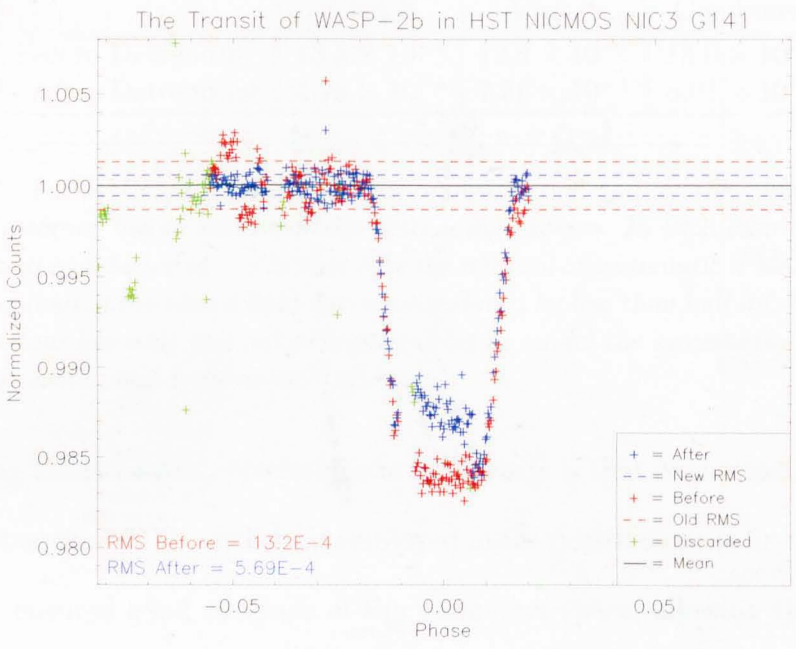


Figure 7.10: *The transit floor in the visit 4 light curve adversely affected by the detrending process. The points before detrending (red) show a transit depth of* $\sim 1.6\%$ *- a figure comparable with the transit depth in visit 3. However, with both temporal parameters (*P_8 *and* P_9*) included in the detrending process and with poor coverage of other parameters (particularly the rotation of the first order spectrum, see Figure 7.8) the transit floor after detrending has become slanted and moved to a depth of* $\sim 1.2\%$*. With systematic effects this poorly constrained, a more reliable model light curve can be obtaned by omiting the detrending step altogether for the transit floor.*

rameters provide inadequate coverage of the transit floor. As a result, the Visit 4 detrended light curve is likely to have been adversely affected by the detrending process (as can clearly be seen in Figure 7.10). Therefore, the in-transit orbit of Visit 4 was omitted from the detrending process and compiled into the final light curve in its raw form (Figure 7.5). This final light curve can be found in Figure 7.11.

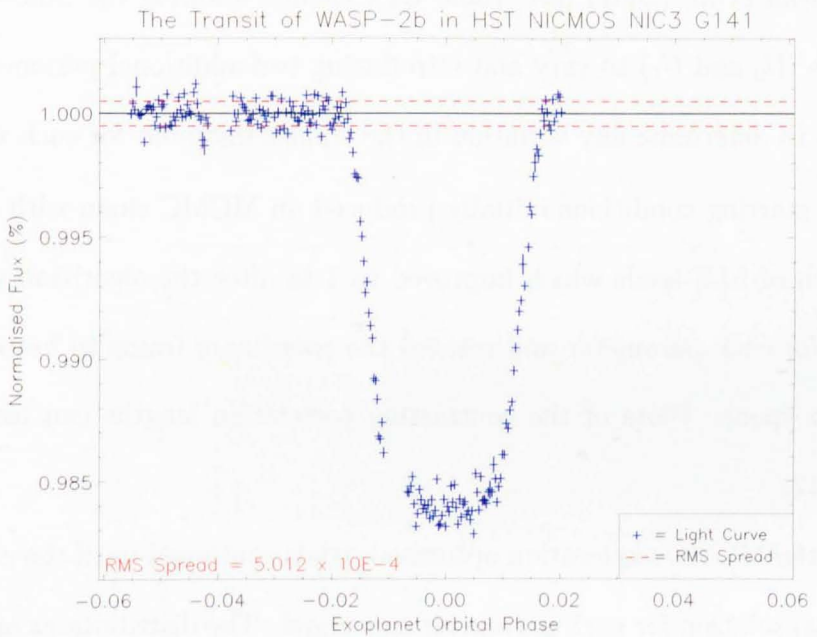


Figure 7.11: *The final, detrended transit light curve for the WASP-2 HST NICMOS NIC3 G141 data set. Only those images whose flux values will be carried forward into further phases of calculation are shown. As in Figure 7.6, the dashed red lines represent the RMS Spread of the Out-Of-Transit points.*

7.5 Light Curve Modeling

The complete light curve was then modelled by application of the MCMC algorithm described in Chapter 5. Initial estimates of the relevant parameters and uncertainties were taken from the results of previous research conducted by Charbonneau et al. (2007) (see Table 7.1); though allowing the limb-darkening coefficents (U_0 and U_1) to vary and introducing two additional parameters (ΔT_3 and ΔT_4) to determine any variation in the transit midpoint for each visit.

These starting conditions initially produced an MCMC chain with a correlation length of 51.6 trials which improved to 1.48 after the algorithm scaled the step size for each parameter and rotated the coordinate frame to better explore parameter space. Plots of the contrasting correlation lengths can be found in Figure 7.12.

With the MCMC exploration optimised, trials continued until the calculation of the ideal solution for each parameter converged. The distributions of accepted trials in the MCMC chain, together with this ideal solution and associated uncertainties, can be found in Figure 7.13.

Note the clear degeneracy between the semi-major axis (a) and inclination angle (i) parameters; resulting in the lenticular distribution along the direction of the degeneracy. This effect is to be expected; as a reduction in both a and i would have the planet following an identical path across the star. However, this does not produce an identical transit light curve since the relative size of the star and planet would also need to change. As a result, the degeneracy is curtailed

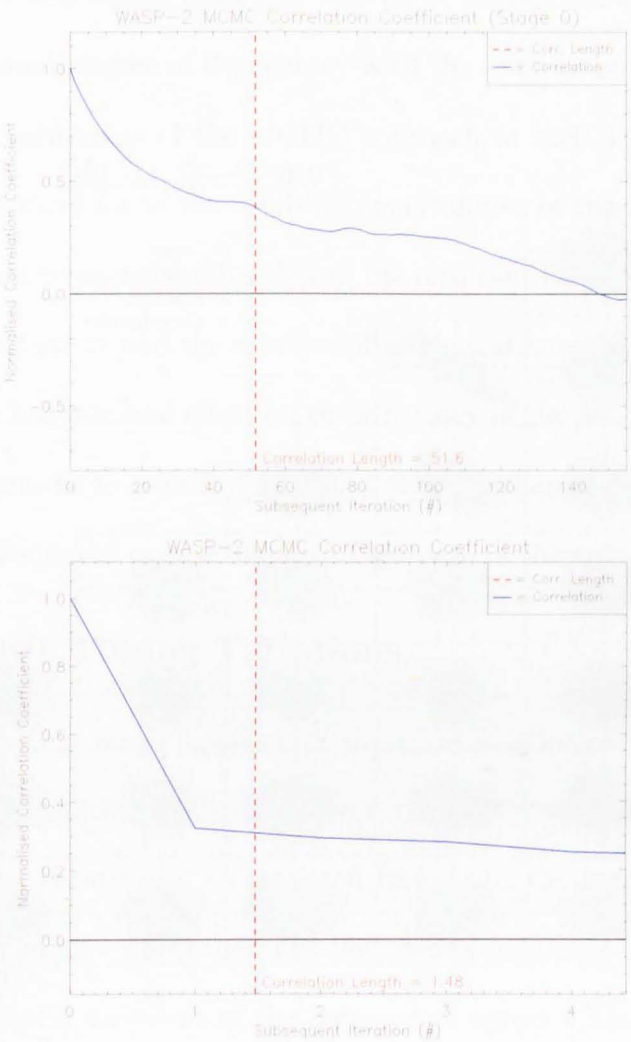


Figure 7.12: *The correlation coefficient calculated along the MCMC chain at the start of processing (top) and once the exploration of parameter space had been optimised (bottom). The inefficient starting exploration exhibits a long trend as each subsequent light curve remains comparable to those which preceeded it (demonstrating that the chain remains in a similar position in parameter space and is not exploring properly). In contrast, the lower curve quickly drops to a background level of correlation where each subsequent light curve in the chain is as different to the starting light curve as any other.*

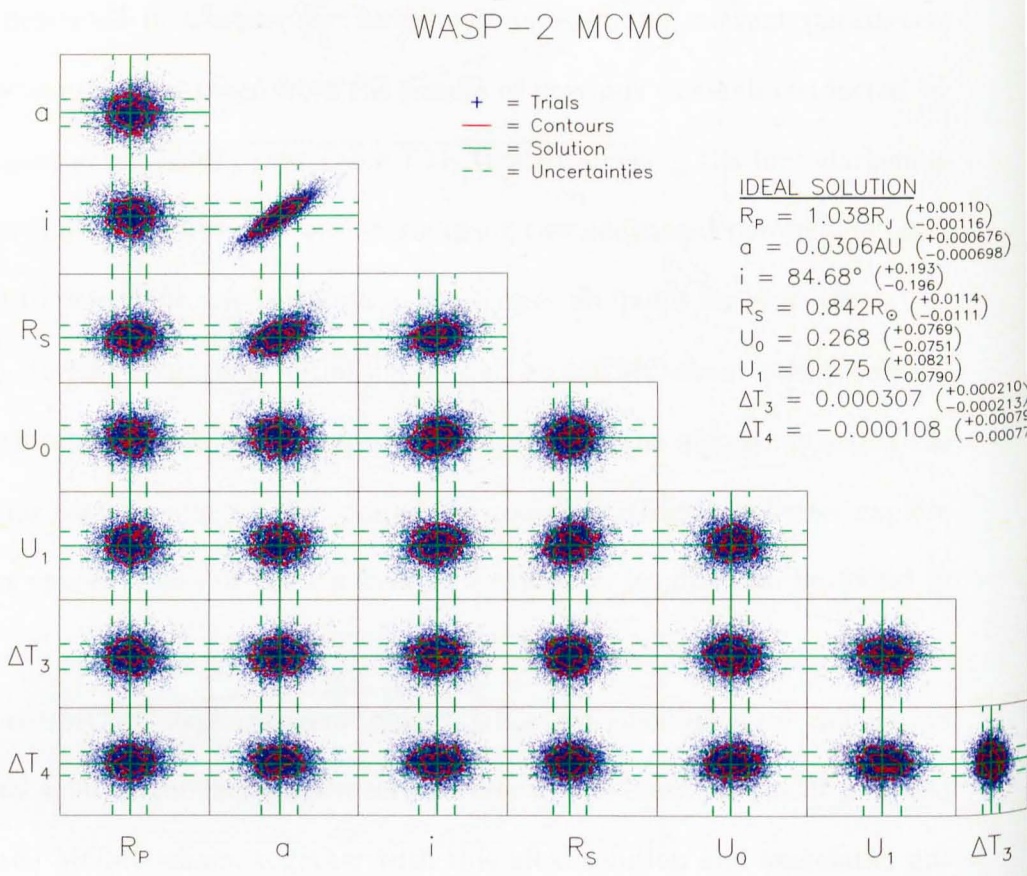


Figure 7.13: *The distribution of accepted trials shown comparatively between all the parameters in the MCMC chain. Circular distributions indicate a uniform, independent parameter. Eliptical or lenticular distributions demonstrate a degeneracy between the two parameters in question. The overplotted solid green lines show the ideal solution derived from this chain, while the dashed green lines demonstrate the uncertainties on that solution.*

by the planetary and stellar radii (R_P and R_S) - though the stellar radius does demonstrate a small degree of degeneracy with the semi-major axis.

An inherent advantage of the MCMC approach to such a fit is that the degeneracy is accounted for in the resulting uncertainties of the ideal solution and therefore does not impact the reliability of the resultant fit. Furthermore, thanks to the addition of priors and the rotation of parameter space (see Chapter 5) the degeneracy also has minimal effect on the efficiency of the MCMC chain.

The histograms fit to each distribution provide perhaps a better visualisation of the ideal solution and can be found in Figures 7.14 through 7.17.

7.5.1 Transit Timing Variations

The most interesting result appears in the transit-midpoint of the third visit (ΔT_3) which is significantly displaced from the expected midpoint (0.0). The parameter is well constrained - calculated from both the ingress and egress of the same transit-event - and is perhaps best demonstrated in Figure 7.18 which demonstrates the displacement of the ingress and egress when the transit is reflected about its uncorrected midpoint.

While an interesting result, it is impossible to say more on such transit timing variations without further measurements. The poor precision of the value of ΔT_4 makes it impossible to perform a useful comparison between the two visits and so more light curves with better temporal resolution would be required to confirm this result. However, the precision of ΔT_4 is sufficient to rule out any inaccuracies in the ephemeris or period of WASP-2b's orbit as the cause of the displacement

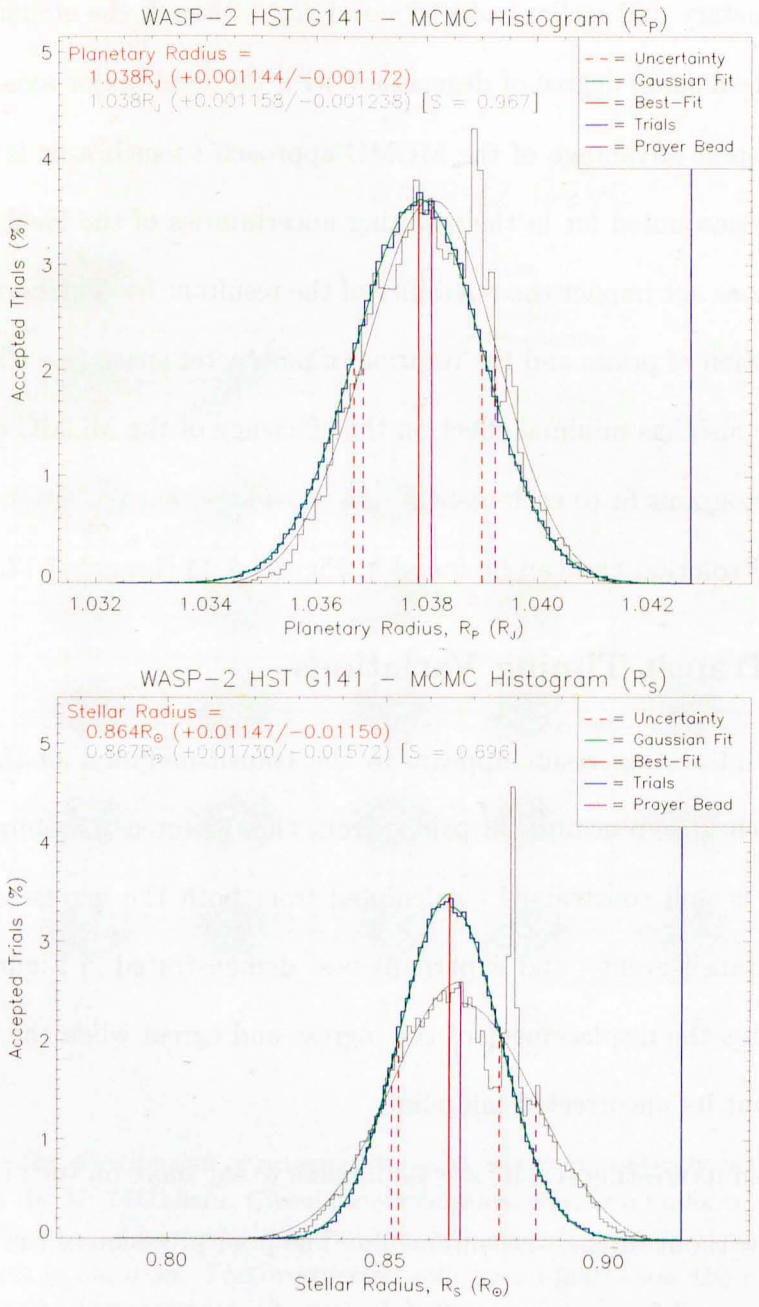


Figure 7.14: *Accepted MCMC trials in the Planetary Radius (R_P, top) and Stellar Radius (R_S, bottom) parameters. The trials have been binned into histograms (blue) and fitted with an asymmentric Gaussian curve (green) from which the ideal solution (solid red line) and related uncertainties (dashed red line) have been derived. The results of Prayer Bead analysis are shown in the grey histogram, with the associated Prayer Bead solution shown in purple.*

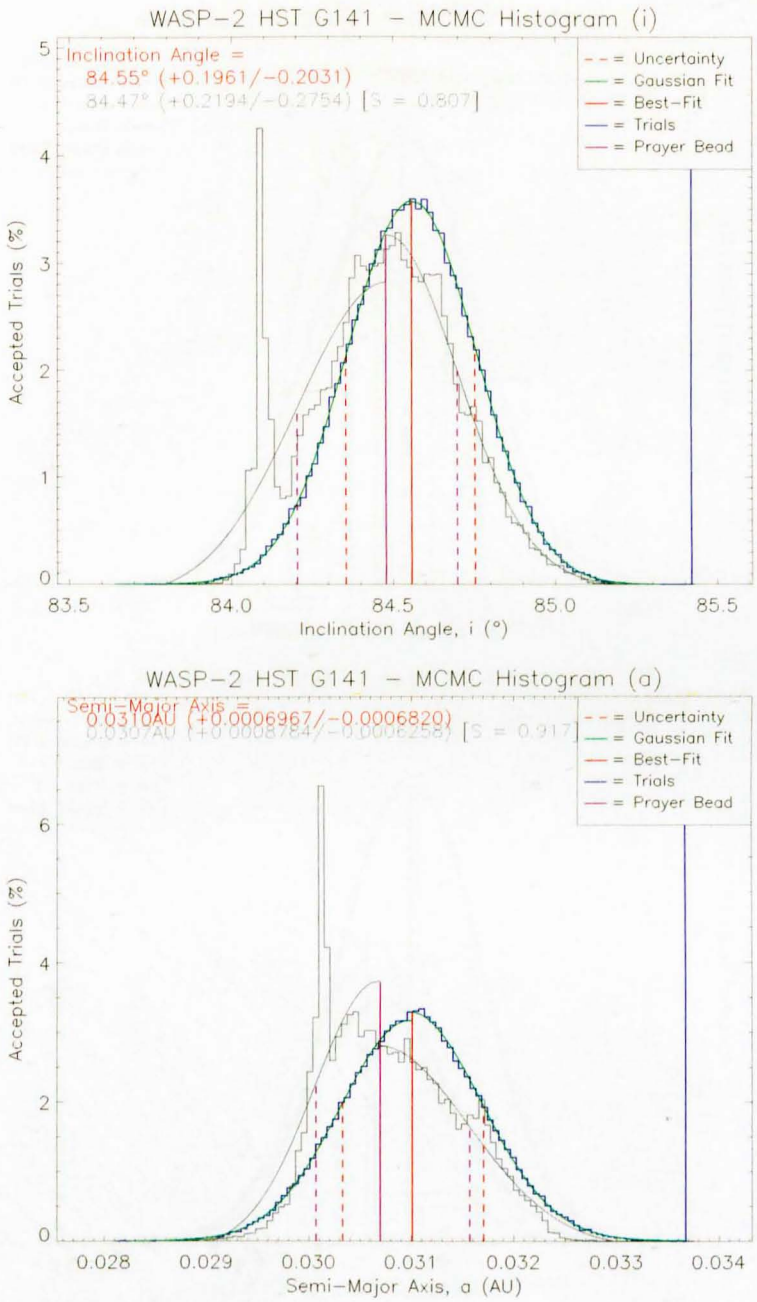


Figure 7.15: *Accepted MCMC trials in the Inclination Angle (i, top) and Semi-Major Axis (a, bottom) parameters in the same presentation as found in Figure 7.14.*

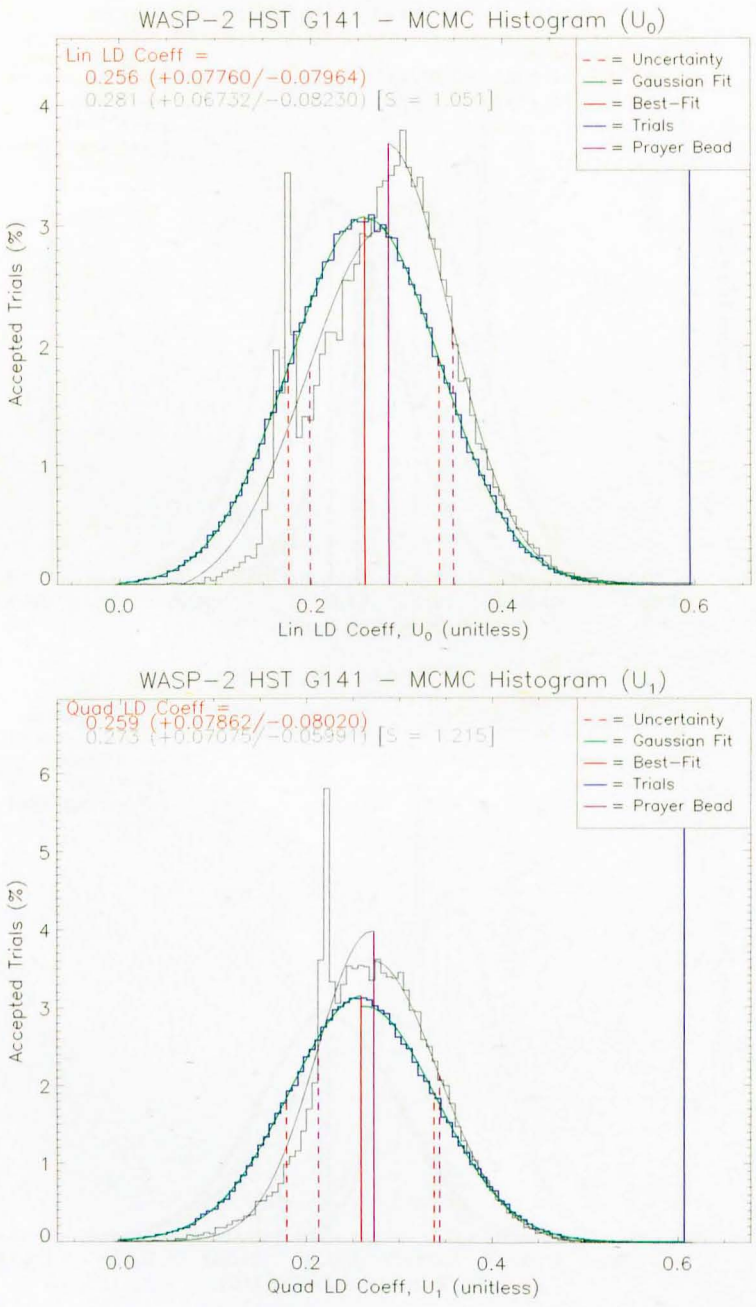


Figure 7.16: *Accepted MCMC trials in the two Limb-Darkening parameters (U_0, top; and U_1, bottom) in the same presentation as found in Figure 7.14.*

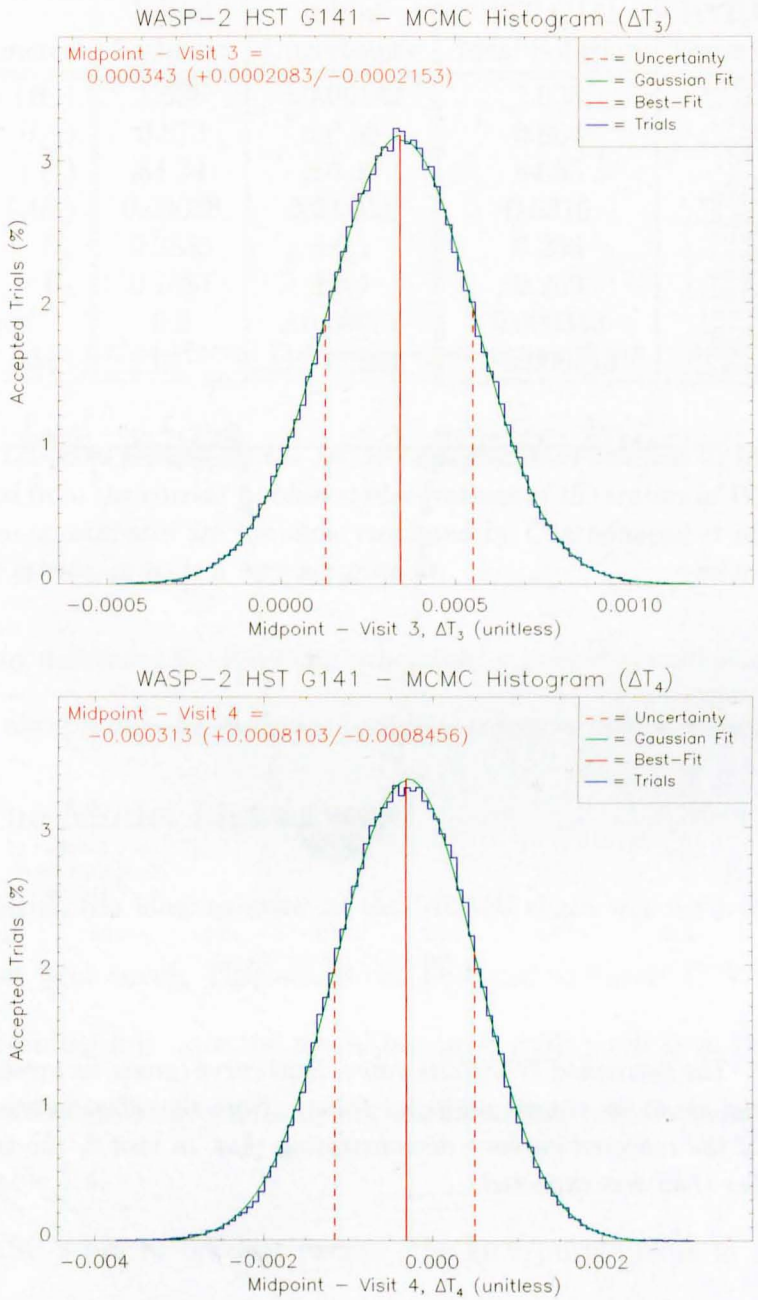


Figure 7.17: Accepted MCMC trials in the Transit Midpoint parameters for visit 3 (ΔT_3, top) and 4 (ΔT_4, bottom) in the same presentation as found in Figure 7.14. Note the far larger uncertainty measurements in visit 4 - caused by that visit having no ingress or egress to help constrain the centre of the transit. Also note the significant displacement of the visit 3 ideal solution from the expected value of 0.

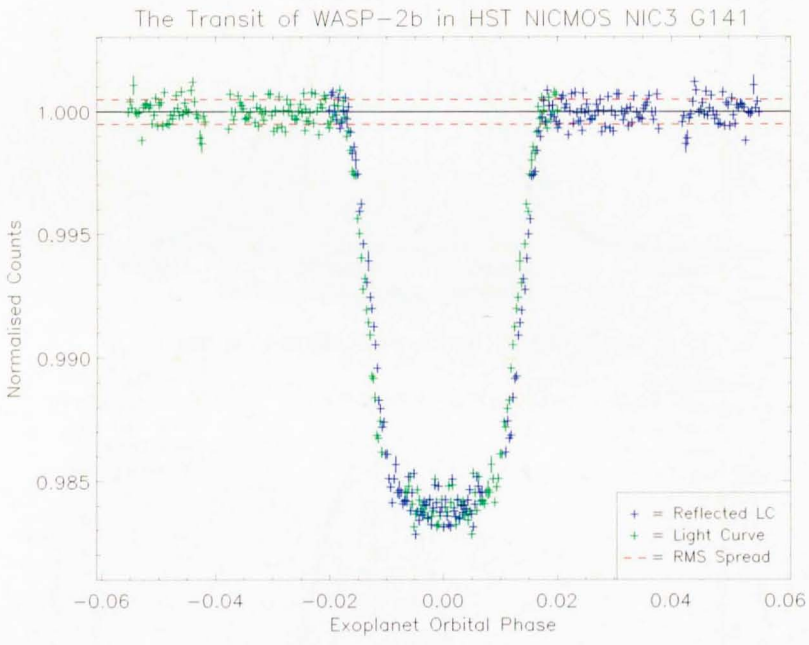


Figure 7.18: *The detrended WASP-2b transit light curve (green, as presented in Figure 7.11) reflected about its transit midpoint (blue). Note the offset between the ingress and egress of the reflected curves - demonstrating that, in visit 3, the transit occured slightly earlier than was expected.*

Parameter	Initial Estimate	Initial Uncertainty	HST G141 Ideal Solution	HST G141 Uncertainty	S Value
R_P (R_J)	1.038	± 0.00142	1.038	$^{+0.001144}_{-0.001172}$	0.967
R_S $(R_\odot)$	0.813	± 0.08	0.864	$^{+0.01147}_{-0.01150}$	0.696
i $(^\circ)$	84.74	± 0.39	84.55	$^{+0.1961}_{-0.2031}$	0.807
a (AU)	0.03038	± 0.0011	0.0310	$^{+0.0006976}_{-0.0006820}$	0.917
U_0	0.2835	± 0.1	0.256	$^{+0.07760}_{-0.07964}$	1.051
U_1	0.2887	± 0.1	0.259	$^{+0.07862}_{-0.08020}$	1.215
ΔT_3 $(Period^{-1})$	0.0	± 0.00001	0.000343	$^{+0.0002083}_{-0.0002153}$	N/A
ΔT_4 $(\Delta\phi)$	0.0	± 0.00001	-0.000313	$^{+0.0008103}_{-0.00084560}$	N/A

Table 7.4: *The ideal solution of the MCMC simulation in contrast to the initial estimates derived from the current published observations of the transit of WASP-2b. The Limb-Darkening estimates are the same employed by Charbonneau et al. (2007) and, as the result attests, provide a very accurate fit.*

in ΔT_3 as any difference in either the ephemeris or period of sufficient magnitude would have also produced a discernable displacement in visit 4, which is not seen.

7.5.2 The Model Light Curve

Once calculated, the ideal solution of the MCMC chain was used to generate a model transit light curve. This model can be found in Figure 7.19 along with a plot of the residual flux once the model has been subtracted from the detrended data. The complete ideal solution, uncertainties and initial estimates are presented in Table 7.4.

The WASP-2 prayer beading results (the grey/purple plots in Figures 7.14 through 7.17) reveal a very different residual systematic situation than was previously expected. While we would anticipate the majority of the residual systematic effects to be concentrated in the in-transit points where the detrending model had to be extrapolated, here we see a distinct feature in the prayer-beading solution.

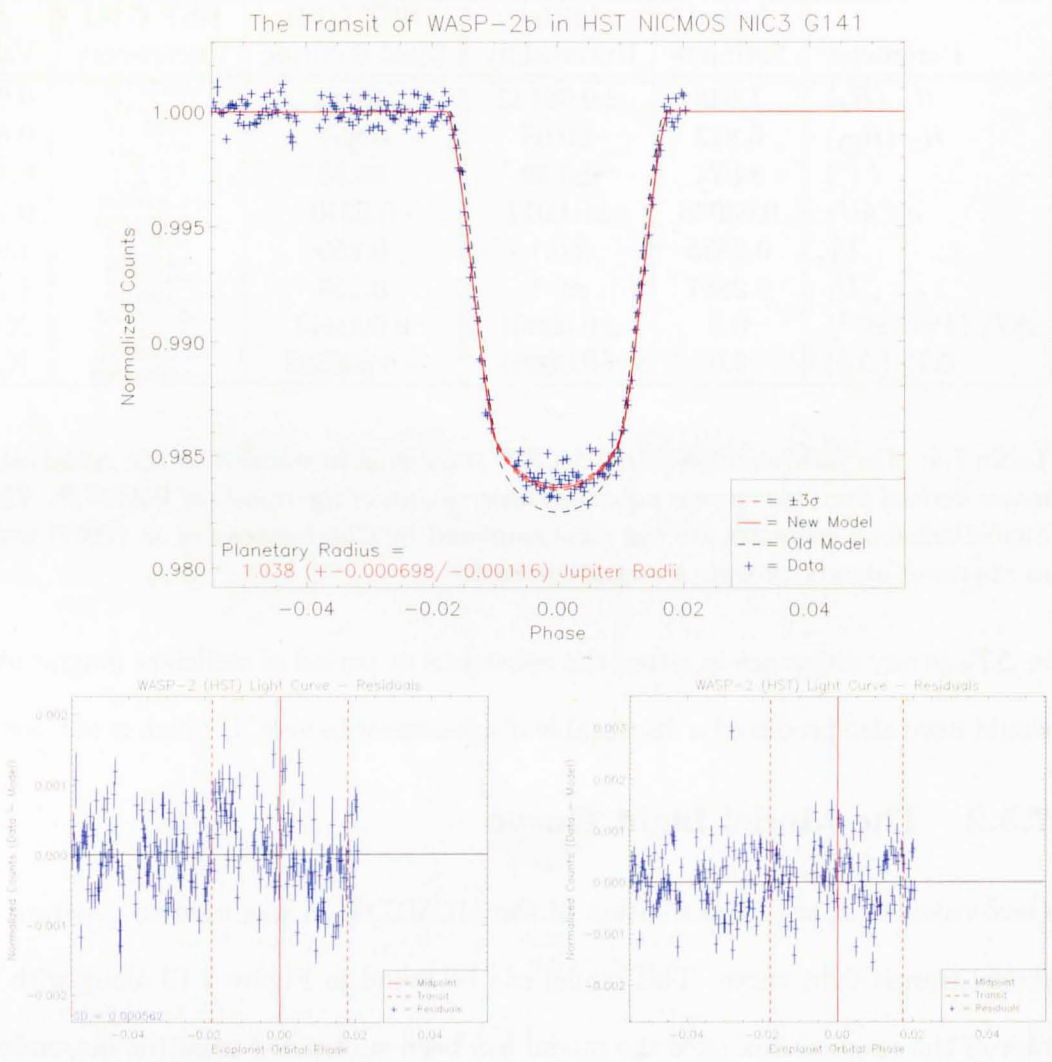


Figure 7.19: *The upper plot shows the detrended transit light curve of WASP-2b (blue, as found in Figure 7.11) overplotted with the model light curve (red) created from the ideal solution of the MCMC simulation. The dashed red lines demonstrate the uncertainties on the model and the dashed black line demonstrates the starting parameters of the MCMC chain. The lower plots demonstrate the residuals when the model light curve is subtracted from the measured light curve with the solid red line representing the transit midpoint and the dashed lines the full duration of the transit. The left-hand plot shows the residuals with no midpoint correction (ΔT_3 and ΔT_4) and a clear displacement can be seen for those points in the ingress and egress. In comparison, the right-hand plot (with ΔT_3 and ΔT_4 applied) shows comparable spread both in and out of transit.*

While the majority of the prayer-beading trials follow the trials of the Ideal Solution, one distinct feature persists in each histogram. Rather than reducing the range of possible solutions (and producing a thinner Gaussian curve) by moving the more believably detrended points to the in-transit parameter space, the S values for the WASP-2 results are typically less than zero - averaging at 0.942 - implying the existence of a large source of systematic noise in the out-of-transit points which is not visible to the naked eye.

The most likely source of this feature is the single out-of-transit orbit for Visit 4. With the transit floor removed from processing, the sole remaining orbit of that visit persists with no secondary orbits to corroborate its detrending model. While the chosen solution may produce a superior RMS spread of the points, there is clearly some residual structure in the orbit which has a large influence on the transit fitting once the prayer beading process moves that region of the light curve into the in-transit parameter space.

Leaving aside the Visit 4 feature in the prayer beading histograms, the rest of the results seem fairly consistent with the remaining light curve (Visit 3). This is confirmation that removing the transit floor of Visit 4 was the correct course of action since the Visit 3 systematic model appears to have left very little residual systematic noise in the final light curve.

7.6 Atmospheric Analysis

As detailed in Chapter 6 the first order spectrum can be binned by wavelength to produce wavelength-dependent light curves and thereby determine any vari-

ation in transit depth with wavelength. Determining the width of the PSF in the WASP-2 data set gives an ideal bin-width of 10.69 pixels (measured from the larger PSF of Visit 4) - covering roughly $0.085\mu m$ in the wavelength domain. Note, however, that since the first order spectrum is defocussed, such a bin actually contains counts from wavelengths spanning double that range. Using 117 such bins, overlapping each by a separation of one pixel, produces wavelength-dependent light curves like that found in Figure 7.20.

As with any NICMOS-derived light curve, the wavelength-dependent curves are detrended prior to the modeling process. It is essential that this step be as robust as possible, since any residual systematics may introduce unreal variations in the transit depth; thereby contaminating the resulting transmission spectrum.

As is clear from Figure 7.20, the detrending of visit 4 (covering the transit floor) would be required to move the depth of transit by a significant degree to bring it in line with the base of the egress in visit 3. As was discussed in Section 7.4, the reliability of the visit 4 detrending is suspect due to its poor parameter coverage and only one out-of-transit orbit against which to model long-term trends. As such, it was considered safer to omit visit 4 from the wavelength dependence analysis altogether and model the transit depth against the ingress and egress of visit 3 alone. With the other planetary parameters constrained by the overall light curve, variations in R_P should still be measurable without the presence of the transit floor. Since the only other parameters expected to vary with wavelength are the limb darkening parameters (U_0 and U_1), the ingress and egress should be sufficient to measure any change in depth with wavelength.

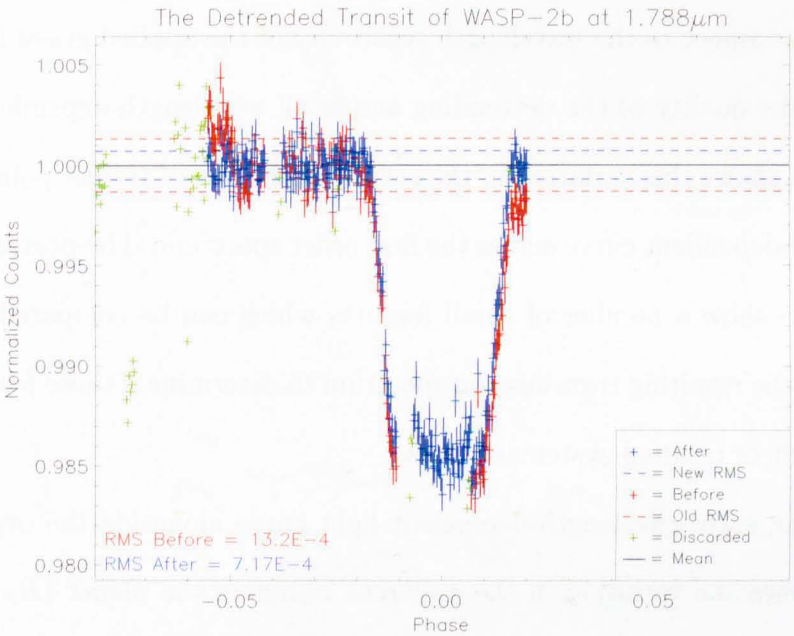


Figure 7.20: *A wavelength dependent light curve from the WASP-2 data set, binned with a central wavelength of* $1.788\mu m$*. The curve is shown both before (red) and after detrending (blue), with points measured from omitted exposures shown in green. Note the clear displacement of the transit floor in relation to the base of the egress - the result of systematics in the visit containing that portion of the light curve. As was noted in Section 7.4, the detrending of visit 4 cannot be considered sound and, as such, detrending has been omited in the case of the transit floor. Even had detrending been performed on the orbit containing the transit floor and even if this process had resulted in an apparent alignment of the visit 4 transit floor and the visit 3 ingress, it would still be impossible to state with confidence that the visit 4 detrending had been reliable, given the problems detailed in Section 7.4. As such, it is best to omit the transit floor from the wavelength-dependent analysis altogether.*

Figures 5.2 and 5.4 demonstrate the impact of changes in these three parameters (R_P, U_0 and U_1) on the transit light curve.

As discussed in Chapter 6, analysing the RMS of the wavelength-dependent curves with respect to the wavelength sensitivity of the applied grism is a means of testing the quality of the detrending across all wavelength-dependent curves. Figure 7.21 shows this variation in the spread of the out-of-transit points in each wavelength-dependent curve across the first order spectrum. The post-detrending RMS values show a number of small features which can be compared with any features in the resulting transmission spectrum to determine if those features may be the result of residual systematic noise.

Modeling each wavelength-dependent light curve alongside the overall transit determines the variation in the apparent radius of the planet (R_P) over the wavelength-range of the first order spectrum (see Figure 7.22). This complete set of wavelength-dependent model light curves is shown in Figure 7.23, demonstrating a clear structure in the variation in transit depth with wavelength. However, as is clear from 7.22, the wavelength-dependent light curves are adversely effected by systematics.

7.6.1 Discussion

As is evident in Figure 7.22, the base of the model transit floor does not match with the base of the transit egress in the data. This is caused by an asymmetry in the ingress and egress, most likely the result of poor detrending due to insufficient coverage of the systematic parameter space, imprecision or extensive systematics

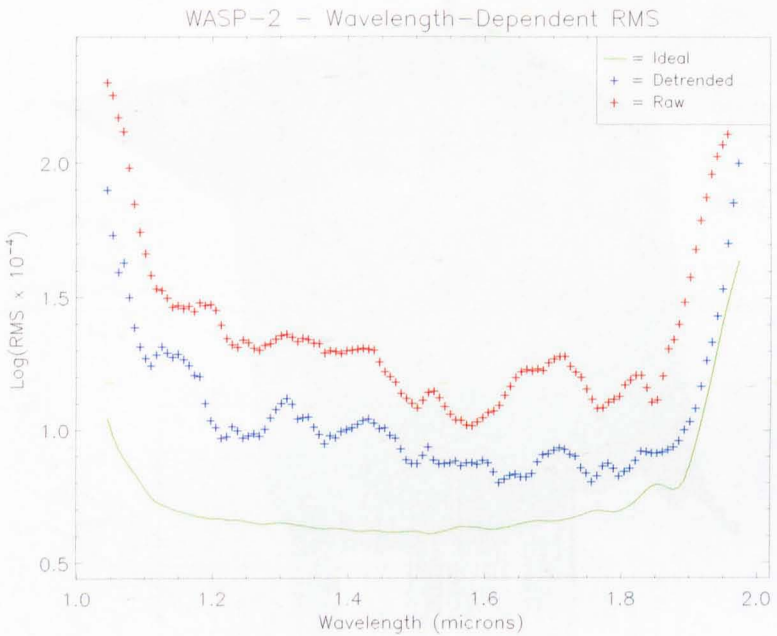


Figure 7.21: *The variation in out of transit RMS Spread across all wavelength dependent transit light curves taken from the WASP-2 data set. The red and blue points show the RMS of each curve both before and after detrending respectively. The overplotted green line represents the theoretical minimum RMS spread achievable (as discussed in Chapter 6).*

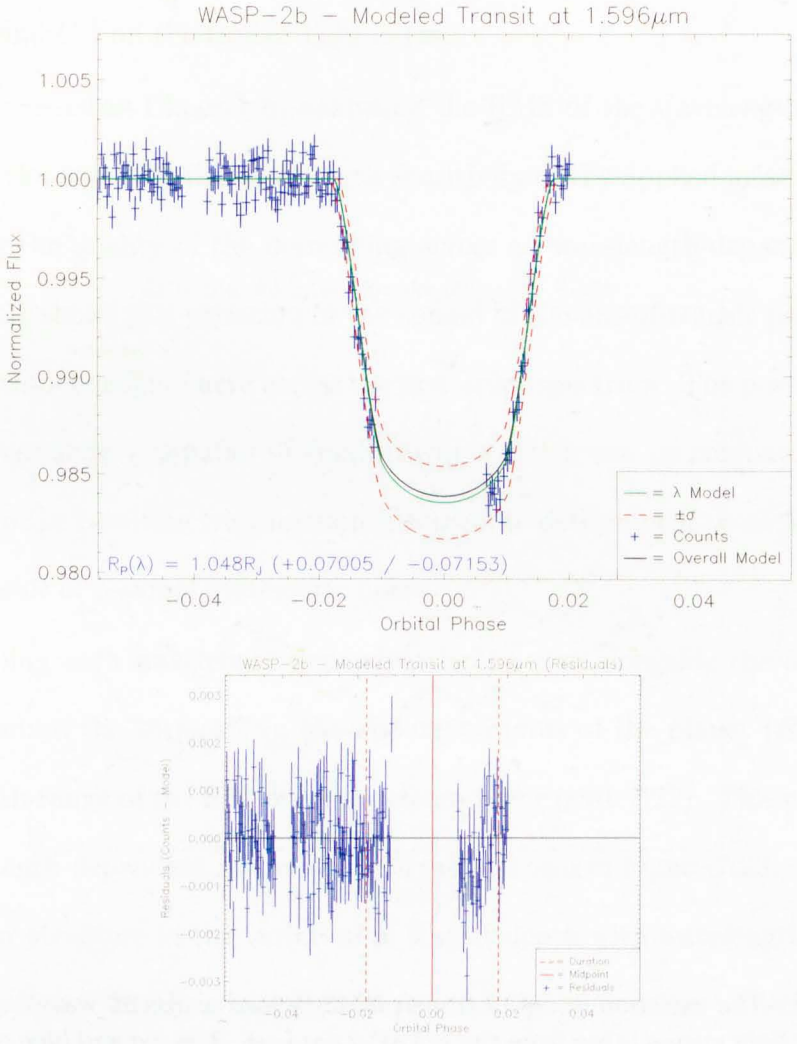


Figure 7.22: *A modelled wavelength dependent WASP-2 transit light curves at a wavelength of* 1.596 *(blue, top) together with the corresponding residuals of the fit (blue, bottom). In the model plot, the green lines represent the model fit to the data and the red lines the uncertainty on that fit. The black line shows the ideal solution of the overall light curve for contrast. The red lnes in the residuals plot show the transit midpoint (solid) and duration (dashed). Residual systematic effects are evident at all wavelengths.*

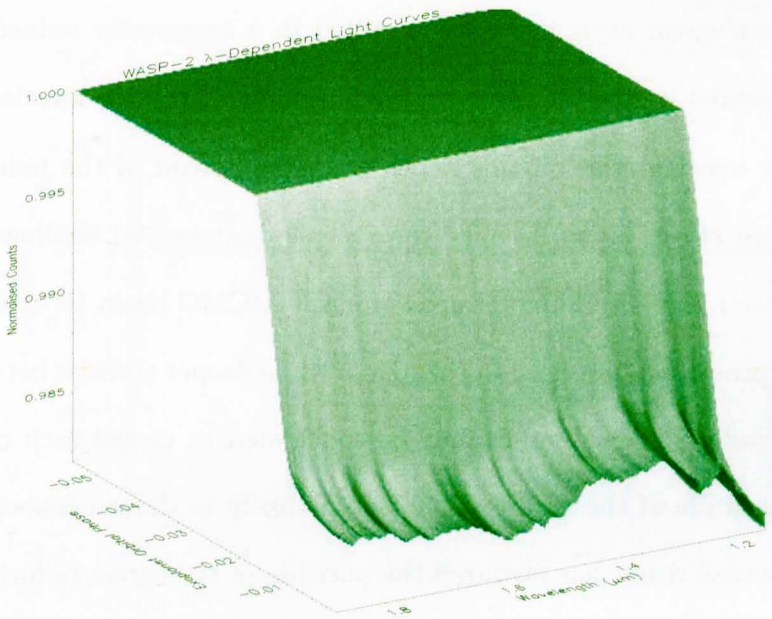


Figure 7.23: *The continuum of wavelength dependent model light curves generated for the WASP-2 data set. While a clear variation of transit depth with wavelength can be observed, it is difficult to draw any firm conclusions due to the residual systematics visible in the data (see Figure 7.22).*

in the detrending parameters themselves (see Chapter 4) or a source of systematic noise not included in our model. From inspection of the light curves and residuals, this systematic appears to have had the same impact across all wavelengths; however, since there is no precise means of quantifying the effect of a systematic in this instance, it is impossible to state this with certainty.

This ingress/egress asymmetry has resulted in a universally reduced transit depth (as measured by the MCMC fitted light curves) across all wavelength bins and far larger uncertainties on this depth than the spread of the points would suggest. These effects arise because the steeper ingress and shallowed egress provide a wider range of parameter space for the MCMC chain to explore; less-deep transits produce a better fit to the egress while deeper transits better match the ingress. Sadly these effects cannot be considered to cancel each other out, since the magnitude of the systematic effects is likely to differ between the two orbits. The overall result has favoured the position of the egress (which has half again as many points as the ingress and so has a greater impact on the χ^2 of any fit) resulting in the overall reduction in the measured ΔR_P across all wavelengths.

Considering just the variations in R_P in isolation, we can calculate the relative absorption of WASP-2b's atmosphere with wavelength - its transmission spectrum. This spectrum is shown in Figure 7.24, though the residual systematics in the wavelength dependent light curves mean any features in this spectrum cannot be believed.

Figure 7.24 provides more evidence of latent systematic errors in the wavelength dependent light curves. On first inspection, the magnitude of the variation

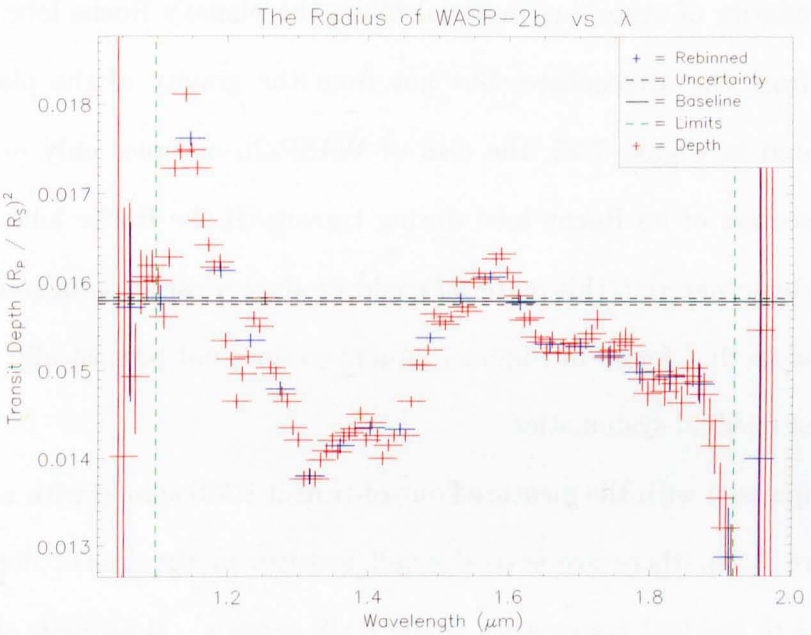


Figure 7.24: $(R_P(\lambda)/R_S)^2$ *for the transit of WASP-2b, measured from a series of wavelength-dependent light curves taken by application of overlapping bins (red). The blue points show the transit depth rebinned to one PSF-width. The overplotted black line shows the depth measured from the overall transit with the uncertainties on that figure represented by the dashed black lines. Residual systematics in the wavelength-dependent light curves have caused the average wavelength-dependent transit depth to be noticably removed from the depth of the overall transit light curve.*

in transit depth appears far larger than would be expected with a total variation of $\sim 25\%$ of the mean transit depth.

While this figure may seem unphysical, this need not be the case. If the atmosphere of WASP-2b is being boiled off by its parent star, we would expect to see a quantity of absorbing material filling the planet's Roche lobe (material liberated from the atmosphere, but not from the gravity of the planet). As demonstrated in Figure 7.25, the disk of WASP-2b occupies only $\sim 11.6\%$ of the cross-section of its Roche lobe during transit. If the Roche lobe was filled with absorbing material, this material could produce a variation in transit depth comparable to that found in Figure 7.24 and so does not intrinsically imply the influence of residual systematics.

In comparison with the measured out-of-transit RMS spread with wavelength (see Figure 7.21), there are several small features in the transit depth which correspond to residual systematics in the RMS spreads[1]. Both plots show small peaks at 1.4 and $1.7\mu m$, indicating that these features are the result of poor detrending rather than WASP-2 atmospheric characteristics. Comparatively, the feature at $1.3\mu m$ in the RMS spread plot corresponds to a trough in Figure 7.24. This is also likely the result of poor detrending, but is most probably caused by a different source of systematic noise than the other features.

Another potential systematic may exist at $\sim 1.15\mu m$ where a feature in the RMS spread curve matches with an apparent step-change in the spectrum where

[1]While such a change in RMS could also represent a region of noiser pixels, such a region is not apparent on inspection of the uncertainty estimates for each pixel provided with every exposure.

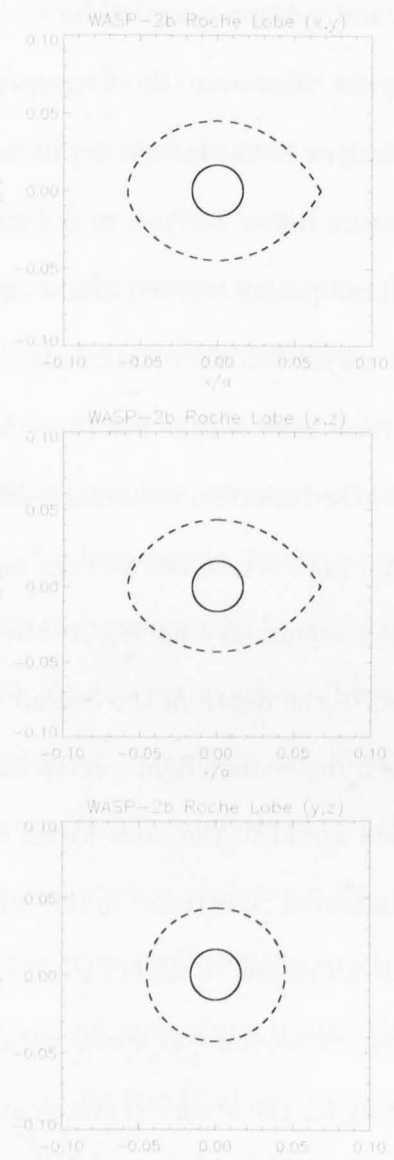


Figure 7.25: *The disk of WASP-2b (solid) shown in comparison with the cross section of its Roche lobe (dashed). The upper plot shows the cross-section in the orbital plane of the system. The middle plot shows the cross-section in the plane perpendicular to the orbital plane and containing our line-of-sight at mid-transit (in both cases, the parent star is to the right of the plot). The lower plot shows the cross section along our line-of-sight at mid-transit - i.e. the cross-section observed during transit. (Busuttil & Haswell, 2011, in prep; and poster presented at NAM2011.)*

the points between 1.13 and 1.17μm seem to be $\sim$ 1% higher than the points on either side of that region. However, the corresponding feature in the RMS spread is wider than the feature in the transit depth, so this may not be the case. It is possible that the feature below 1.17μm in the spectrum is a real incline in the transit depth while those points below 1.13μm are imprecise; due to having been measured from light curves too close to the edge of the first order spectrum. Note that all of these features have widths less than that of the PSF (0.11μm) as would be expected for localised residual systematic effects.

Perhaps most damning, however, is the average measured depth of the transit; with the ingress/egress asymmetry having resulted in a universally reduced $(R_P/R_S)^2$ when compared to the depth of the overall transit. If the residual systematics in the wavelength dependent light curves have the same impact at all wavelengths, this reduction would be the same for all light curves - allowing us to simply renormalise the measured 'spectrum' in line with the depth of the overall transit. However, there is no means to determine if this is the case, or whether the cause of the ingress/egress asymmetry varies with λ.

One means of accounting for the observed systematics would be to reflect each wavelength-dependent light curve about its midpoint and model composite light curves of both the original and reflected transit. This approach would remove the dominance of the egress, but the systematics themselves would still be present and would likely still influence the results.

A similar possibility would be to model the ingress and egress independently - creating two spectra whose combined average depth would be expected to match

that of the overall light curve. While systematics would still be present in both results, the systematic causing the asymmetry is observed to have the opposite effect on each orbit. As a result, we would expect any resulting systematic features in one spectrum to be inverted in its twin - with a peak in the ingress-spectrum corresponding to a trough in the egress spectrum. In contrast, any common features in both spectra could be considered evidence of atmospheric effects[1].

Perhaps a more simple approach to this issue is to apply a more exacting (if less robust) means of modeling the first order spectrum. In comparison with the MCMC process, a simple χ^2-minimization 'bootstrapping' routine will locate on the light curve solution with the minimum χ^2 fit to the data. In this instance, such a fit would be expected to locate on the egress (because of its greater number of points) and ignore the egress. In effect, we would expect a bootstrapping routine to produce results comparable to those of an MCMC chain simulating the egress in isolation. Performing such a process on the rebinned, non-overlapping wavelength-dependent light curves produces the spectrum found in Figure 7.26.

While Figure 7.26 retains an equivalent transit-depth offset, like that seen in Figure 7.24, the features in its spectrum are considerably more plausible. In particular, the features at $< 1.2\mu m$, $1.4 - 1.6\mu m$ and $> 1.75\mu m$ are consistent with absorption by water; as would be expected in such a scenario. In addition to the apparent absorption features, the spectrum also demonstrates an overall gradient which may be consistent with an atmospheric haze in the atmosphere of

[1]Such features could also result from other, unrelated systematic effects currently obscured by the dominant ingress/egress asymmetry. However, once revealed, such systematics (if any exist) could potentially be identified and discounted by comparison with the wavelength dependent OOT RMS curve (Figure 7.21).

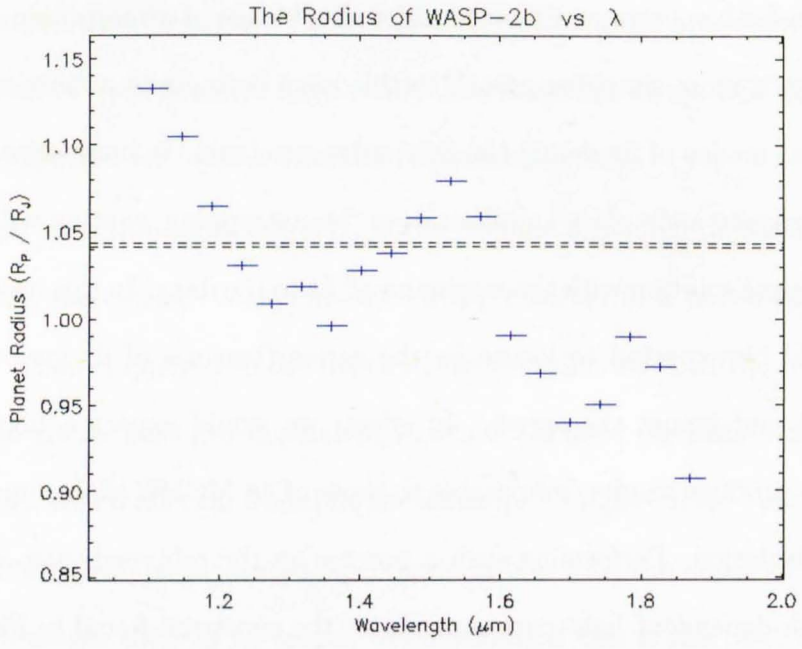


Figure 7.26: *The variation in radius of WASP-2b with wavelength (R_P, blue) measured by a series of bootstrapped models to wavelength dependent light curves. The overplotted dashed black lines show the uncertainty range of the value of R_P calculated in for the overall model light curve.*

WASP-2b itself (rather than the material filling the Roche lobe. However, with the clear systematics in the wavelength dependent light curves, it is currently impossible to draw any firm conclusions from this data.

7.7 Conclusions

The WASP-2 NICMOS G141 data set has yielded the most precise measurements of the system's planetary parameters to-date. However, further analysis of the data - particularly of the transit's wavelength dependence - has been prevented by the prevalence of systematic noise in the data set.

Valuable lessons can be learned from the WASP-2 results to help future studies. Firstly, the lack of a second out-of-transit orbit in the visit four light curve led to a poorly constrained and unreliable detrending process. From examination of the parameters and detrending results (see Section 7.4) we can suggest that an optimum observing strategy is one with at least two wholly out-of-transit orbits and, preferably, in which every orbit has some out-of-transit points to reduce the need to extrapolate across the in-transit parameter space (though depending on the duration of the transit, this may not be possible while also capturing the transit floor).

It should also be noted that the dithering employed in the WASP-2 out-of-transit observations seems to result in a much more reliable detrending of the spatial (y) position parameter, P_0. While such an approach is not feasible for all parameters, it suggests that planned sub-pixel variations in the position of the first order spectrum can guarantee a good coverage of the in-transit parameter

space and so improve the believability of the resulting light curve. This could be acheived by designing a set of observations to dither in P_0 and both the spectral (x) positon, P_1, and rotation, P_5, parameters. It may also be possible to acheive a similar effect in the spatial width and spectral length parameters (P_2 and P_3) by varying the level of defocussing across a selection of out-of-transit exposures.

Despite the possible detection of atmospheric haze and molecular absorption features, the wavelength dependence analysis of the WASP-2 data set serves best as a cautionary tale for the potential impact of systematic noise. While an improved algorithm or better measurements of the sytematic parameters may yet improve the wavelength-dependent light curves, we cannot rule out the possibility that any (or, indeed, all) features in the wavelength-dependence of the transit depth (see Figures 7.24 and 7.26) are the result of systematic noise.

If this is the case then the large scale variations in transit depth are caused by relatively small variations in the light curves from which those figures are derived. It is likely that atmospheric features of a more believable magnitude (not taking into account the possibility of a cloud of absorbing material filling WASP-2's Roche lobe) would correspond to systematic effects too small to be discerned through inspection of the wavelength dependent curves.

Nevertheless, the results of the WASP-2 data set are worthy of further research to determine the true nature of the hot Jupiter's atmosphere. Regardless of the findings reported in Section 7.6.1 are proven or disproven as a result, either scenario will result in a deeper understanding of the systematic noise present in such NICMOS grism observations.

Chapter 8

Results - HD189733

8.1 Chapter Outline

The following chapter details the results of the HD189733b transit data set as derived by application of the data processing pipeline outlined in Chapters 3-4. The chapter begins with a review of the extant research and properties of the HD189733 system - highlighting the current debate in the field as to the reported presence of Methane in the hot Jupiter's atmosphere (a result initially derived from this very data set, Swain et al. (2008)). This is followed by details and analysis of the outputs of the pipeline from the raw light curve, through the removal of systematic noise, to the MCMC modeling of the processed light curve. Finally, the chapter reports the findings of the wavelength analysis of the HD189733b transit - paying special note to whether there is evidence of methane in the resulting transmission spectrum and comparing our results to those of previous analyses of this data set.

Parameter	Value	Uncertainty
Planetary Radius, R_P (R_J)	1.178	$\pm^{0.016}_{0.023}$
Planetary Mass, M_P (M_J)	1.138	± 0.025
Inclination Angle, i (°)	85.51	$\pm^{1.00}_{0.05}$
Semi-Major Axis, a (AU)	0.03142	± 0.00052
Period, P ($days$)	2.21857312	$\pm^{3.6\times10^{-7}}_{7.6\times10^{-7}}$
Transit Epoch, E (HJD)	2453988.80336	± 0.00024
Stellar Radius, R_S ($R_\odot$)	0.788	± 0.051

Table 8.1: *The latest parameters of the HD189733 exoplanetary system, at time of writing, gathered from research conducted by Baines et al. (2009), Southworth (2010) and Triaud et al. (2010).*

8.2 Introduction

HD189733 was found to host a Jupiter-mass exoplanet in 2005 (Bouchy et al., 2005) and was quickly recognised as a prime candidate for atmospheric studies. At time of writing, only HD209458 (which hosts the first known transiting exoplanet) rivals HD189733 as the most studied of this population. The most up-to-date planetary parameters for HD189733b can be found in Table 8.1.

As Figure 8.1 demonstrates, HD189733b is comparable to WASP-2b in its place in the overall transiting exoplanet population. However, while HD189733b is of comparable radius (R_P), its mass (M_P) is significantly greater. While HD189733b is slightly cooler than WASP-2b, this temperature difference is not sufficient to explain this difference in density - suggesting that the cause may be a fundamental difference in the structure or chemical composition of the two planets.

The atmosphere of HD189733b has become the subject of much debate. Ini-

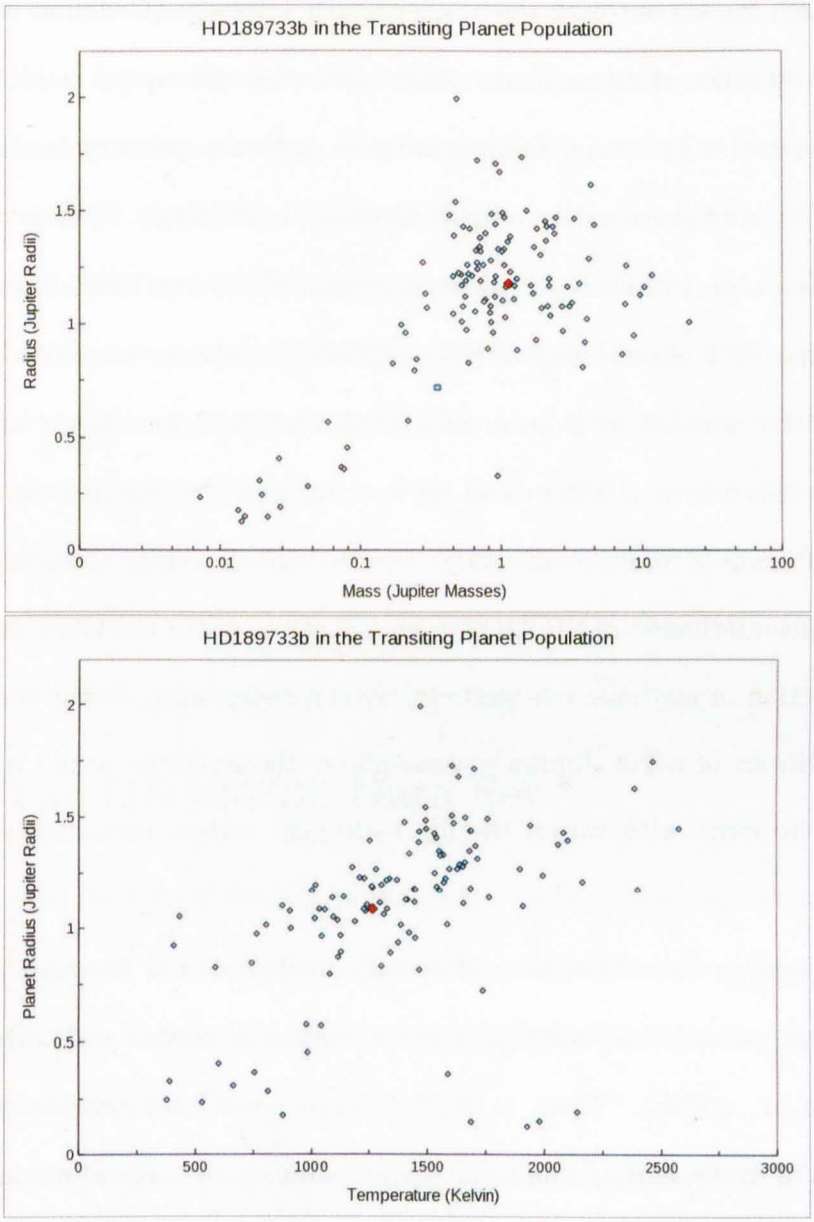


Figure 8.1: *The exoplanet HD189733b (red) in contrast to the population of transiting planets as a whole (blue). The upper plot shows the relation between mass and radius of the planets while the lower plot shows the effect of temperature on the planets' radii. Note that in both cases, HD189733 lies towards the centre of the distribution.*

tial research carried out with the Spitzer Space Telescope (Grillmair et al., 2007) found no evidence of atmospheric water. However, subsequent research contradicted the initial detection; finding water in both the primary transit (Tinetti et al., 2007) and the secondary eclipse (Grillmair et al., 2008). Shortly afterwards, ground-base observations reported the detection of sodium (Redfield et al., 2008).

The first HST observations of HD189733b's atmosphere were also reported in 2008. Initial research by Pont et al. (2008) with ACS found only atmospheric haze between 0.58 and $1.2\mu m$ and no spectral line features of note. However, shortly afterwards Swain et al. (2008) reported the detection of water and, more surprisingly, methane via NICMOS grism spectroscopy in the G206 grism range. The detection of methane was particularly surprising since, in the temperatures and conditions of a hot Jupiter's atmosphere, the molecule would normally be expected to react with water; leaving hydrogen, carbon monoxide and carbon dioxide.

Furthermore, from that same data set, analysis of the dayside spectrum of HD189733b yielded the detection of water, carbon monoxide and carbon dioxide (Swain et al., 2009b). Then, in 2010, ground based observations claimed the detection of water, carbon monoxide, carbon dioxide and, again, methane (Swain et al., 2010).

However, these detections remain in doubt as further research by other groups has repeatedly failed to duplicate the detections. A second examination of the Swain et al. (2008) data by Gibson et al. (2011) found it impossible to conclude that the features in the reported spectrum were not the result of residual sys-

tematic noise (see Chapter 4). Further observations by Sing et al. (2009), again using NICMOS but performed with photometric filters rather than by grism spectroscopy, found the same haze as Pont et al. (2008), but in the same wavelength range as Swain et al. (2008). The ground-based detections (Swain et al., 2010) were contradicted by similar research conducted by Mandell et al. (2011), who suggest that the Swain et al. (2010) detection is likely the result of telluric effects.

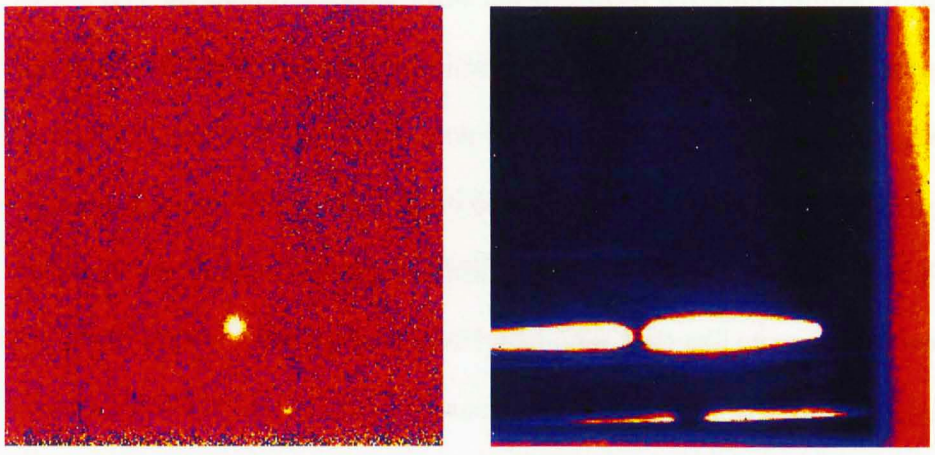
Alongside this controversy, other studies have reported the detection of an atmospheric haze - like that found by Pont et al. (2008) - in the infrared (Désert et al. (2009) with Spitzer and ground-based observations by Barnes et al. (2010)) and optical/near-UV (Sing et al., 2011). The NICMOS observations reported by Swain et al. (2008) are re-analysed in this chapter.

8.3 The HD189733 Data Set

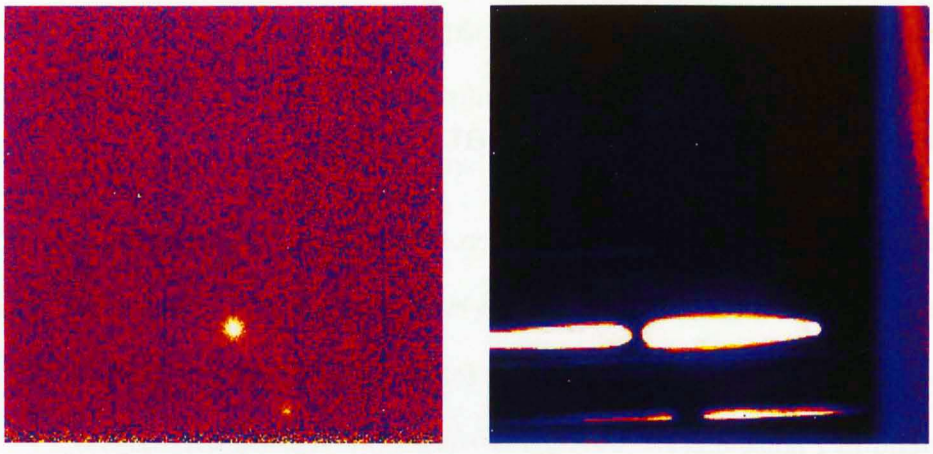
HST observed the HD189733 system across two visits - one to capture the primary transit, the second to attempt an observation of the planet's secondary eclipse. Full details of the observations can be found in Table 8.2 and sample images can be found in Figure 8.2.

Processing the direct images for each visit yielded the position of the star as demonstrated in Figure 8.3 and found in Table 8.2. From this and the grism images, the light gathering box was defined with dimensions shown in Table 8.2 and Figure 8.4.

Counting the light inside the light gathering box for both visits produces the light curves in Figures 8.5 and 8.6. In Figure 8.5, the primary transit is clearly



(a) Visit 1 - Secondary Eclipse

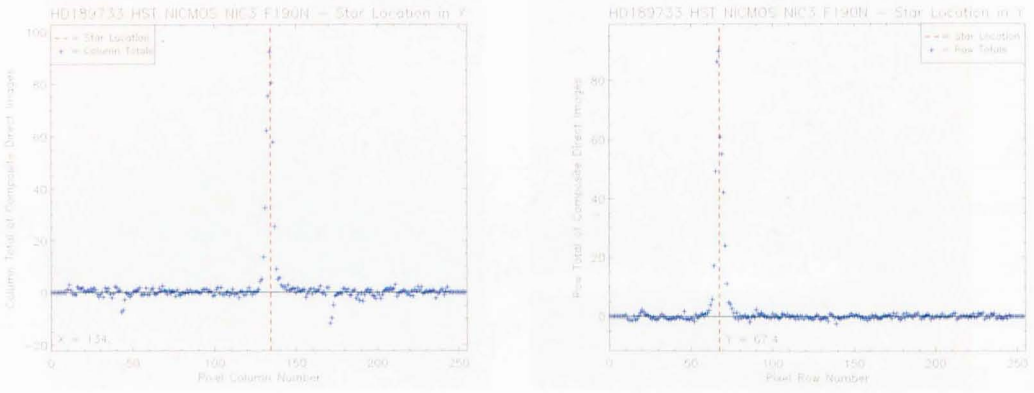


(b) Visit 2 - Primary Transit

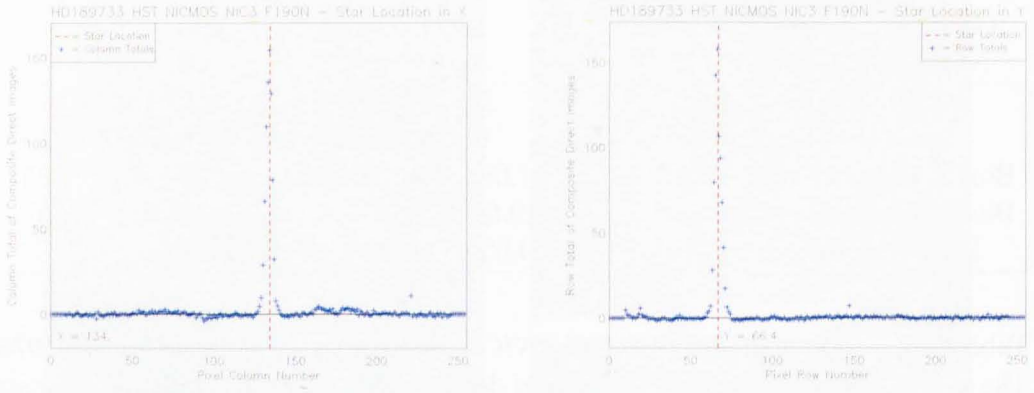
Figure 8.2: *Sample Filter (left) and Grism (right) images from the two visits of observation comprising the HD189733 data set. The smaller object towards the bottom of each image is a background star.*

	Visit 1	Visit 2
Containing	Secondary Transit Floor	Primary Transit Floor
Date of Observations	April 2007	May 2007
Number of Orbits	5	5
Filter	F190N	F190N
Filter Exposure Time (s)	0.406	0.406
Number of Filter Images	5	5
Grism	G206	G206
Grism Exposure Time (s)	1.624	1.624
Number of Grism Images	638	638
Star X Pos ($pixels$)	134.48	134.23
Star Y Pos ($pixels$)	67.46	66.50
Box X Dimensions ($pixels$)	81.08 : 198.20	80.83 : 197.95
Box Y Dimensions ($pixels$)	49.68 : 78.97	49.63 : 77.76
PSF Width ($pixels/\mu m$)	9.63 / 0.111	9.55 / 0.110

Table 8.2: *Details of the HST NICMOS NIC3 observations of the transit of HD189733b. The upper three sections of the table hold details of the exposures and observing strategy, the lower section holds details calculated from the resulting images (as detailed in Chapter 3).*

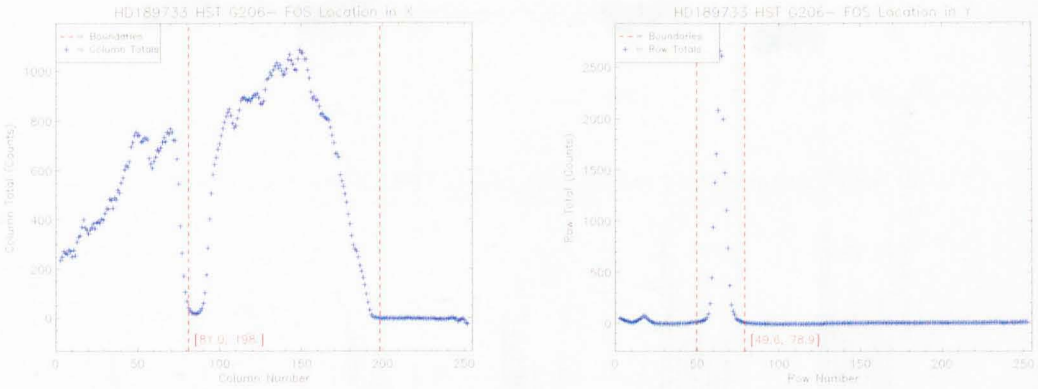


(a) Visit 1 - Secondary Eclipse

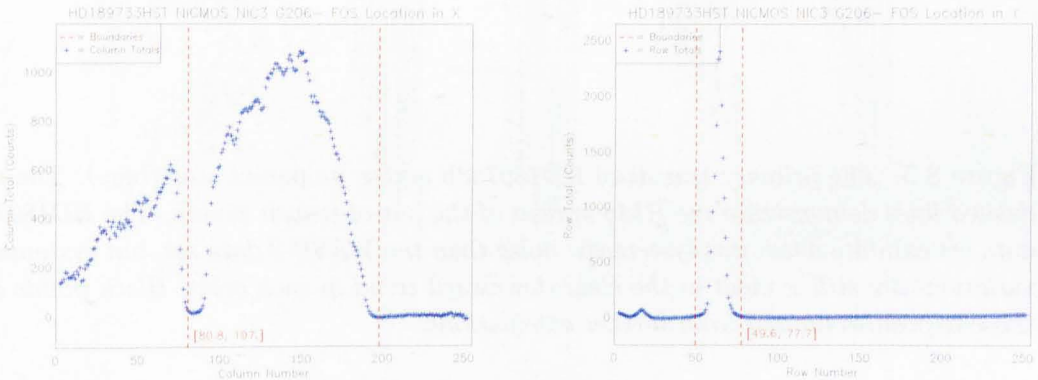


(b) Visit 2 - Primary Transit

Figure 8.3: *The position of the star (red) in the HD189733b data set determined from the column- (right) and row-totals (left, blue) of composite filter images from each visit.*



(a) Visit 1 - Secondary Eclipse



(b) Visit 2 - Primary Transit

Figure 8.4: *The dimensions of the light gathering box shown in both the spectral (left) and spatial axes (right). The blue points are the column- and row-totals of a composite image comprising every grism image in each visit from which the limits of the box (denoted by the red dashed lines) have been calculated (see Chapter 3). Unlike the G141 grism used in the WASP-2 observations, the G206 grism has a second order spectrum of comparable size to the first order spectrum - as evidenced in the left-hand plots. However, as the second order spectrum lies only partially on the detector, it would be detrimental to include its counts in the overall light curve.*

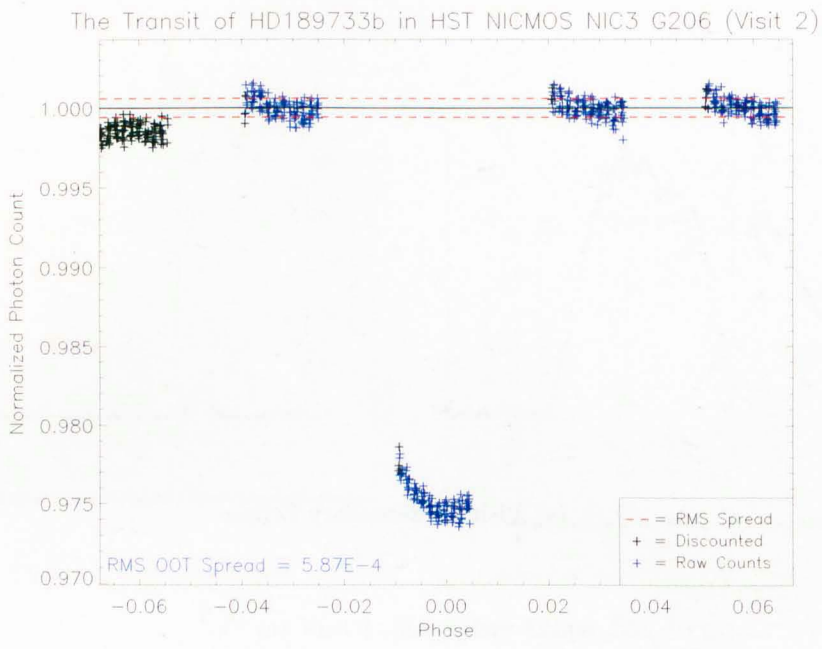


Figure 8.5: *The primary transit of HD189733b across its parent star (blue). The red dashed lines demonstrate the RMS spread of the out-of-transit points. The HD189733 data set exhibits much less systematic noise than the WASP-2 data set, but systematic variations are still evident in the clear downward trend in each orbit. Black points are those exposures omited from further calculations.*

discernable in the third orbit. Furthermore, while systematic effects are visible in the out-of-transit points, their effect appears to be minimal aside from an apparent downward trend throughout each individual orbit.

In contrast, Figure 8.6 has a clear orbit-to-orbit upward trend. The secondary eclipse, somewhat discernable in the third orbit, is noticably displaced below this trend, but it is clear that such systematics must be removed before any firm conclusions can be drawn. Despite how the two curves appear (with the secondary eclipse having a much smaller range of fluxes) the primary transit exhibits the inferior out-of-transit RMS spread.

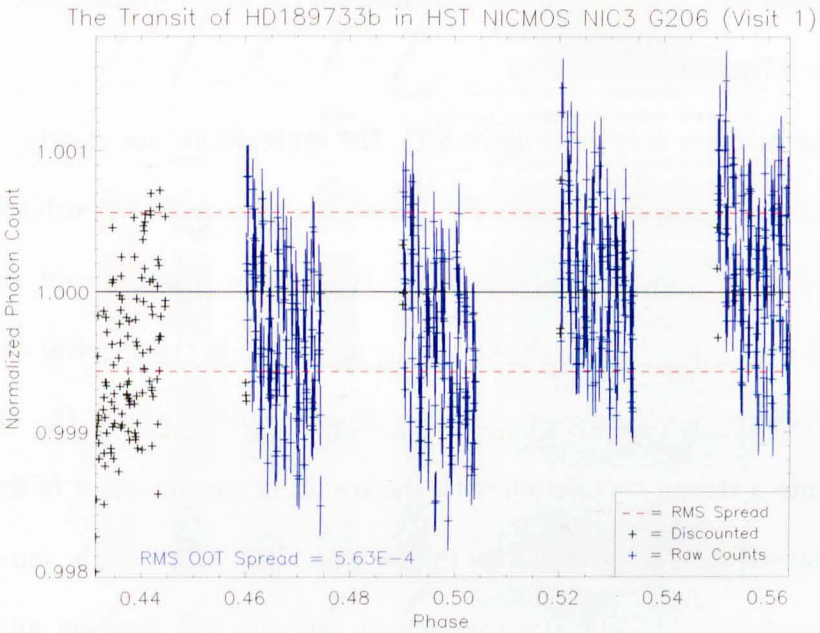


Figure 8.6: *The secondary eclipse of HD189733b by its parent star (blue). In contrast with the primary transit, the secondary eclipse exhibits a consecutive orbit-on-orbit systematic trend rather than the isolated trend in each individual orbit of Figure 8.5. Despite this, the secondary eclipse is discernable in the third orbit which has noticably fewer counts when compared to its fellows. While not directly evident because of the relative scales between the visits, the secondary eclipse light curve has a superior RMS spread (denoted by the dashed red lines) than the primary transit. As before, the black points are those images removed from future calculations.*

8.4 Detrending

After extracting the light curve of each visit, the systematic parameters were measured for each visit and used to model and remove the systematic effects. These parameters, and their modelled effect on the out-of-transit flux are shown in Figures 8.7 and 8.8.

For the primary transit (Figure 8.7), the systematics are clearly dominated by the orbital phase parameter, P_9, which matches with the orbital gradient noted in Figure 8.5. In addition, there is evidence of significant systematics dependent on the position of the first order spectrum in the spectral axis, P_1; the detector background, P_6; the spectral, P_4; and the spatial width, P_2 - all of which demonstrate a strong correlation with the resulting measurement of flux.

Comparatively, the systematics in the secondary eclipse light curve (Figure 8.8) demonstrate the same dependency on the spectral position and detector background (P_1 and P_6), but also exhibit a strong dependence on the spatial position, P_0. Note that a strong correlation is found with the bias parameter (P_{10}) which cannot be safely included in the detrending process due to the offset with the orbit containing the secondary eclipse (see upper-left panel, Figure 8.8).

The apparent groupings of the points in both light curves is a result of the exposure increments employed for the images across each orbit. As can be seen in the HST phase parameter in both visits (P_9 in Figures 8.7 and 8.8), the start-time of each exposure shows several steps as the detector pauses after every twelfth image.

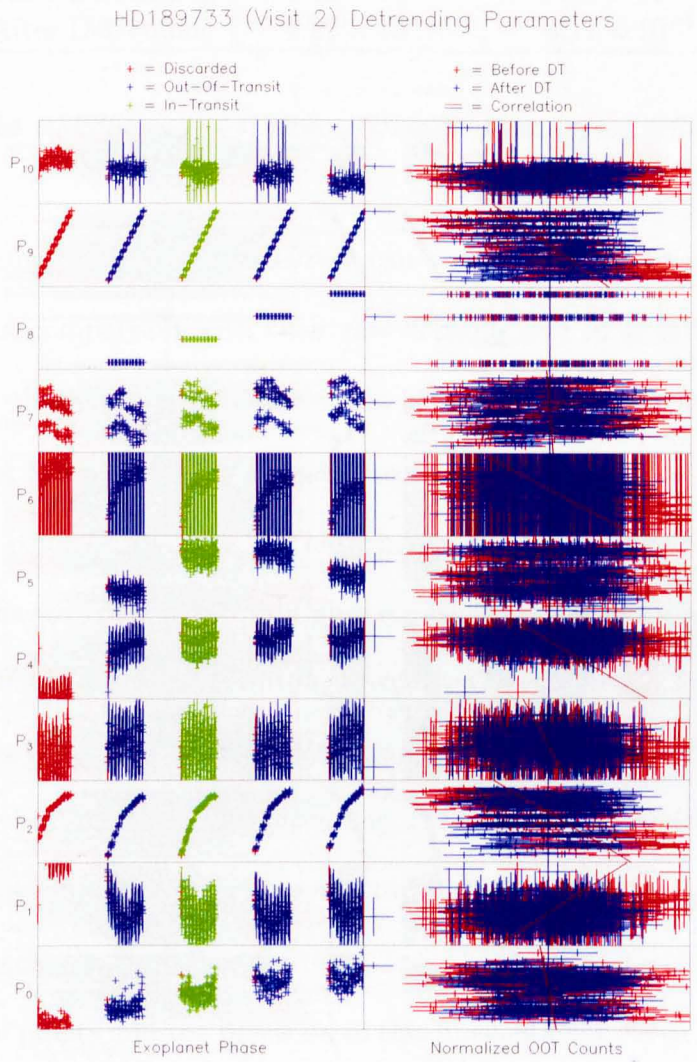


Figure 8.7: *The detrending parameters applied to the primary transit light curve (left) and their modelled correlation with the out-of-transit flux (right). Points on the parameter plots have been colour-coded to distinguish between out-of-transit exposures (blue), in-transit exposures (green) and omitted exposures (red). The out-of-transit correltion plots show the impact of the parameters on the measured flux both before (red) and after (blue) the detrending process.*

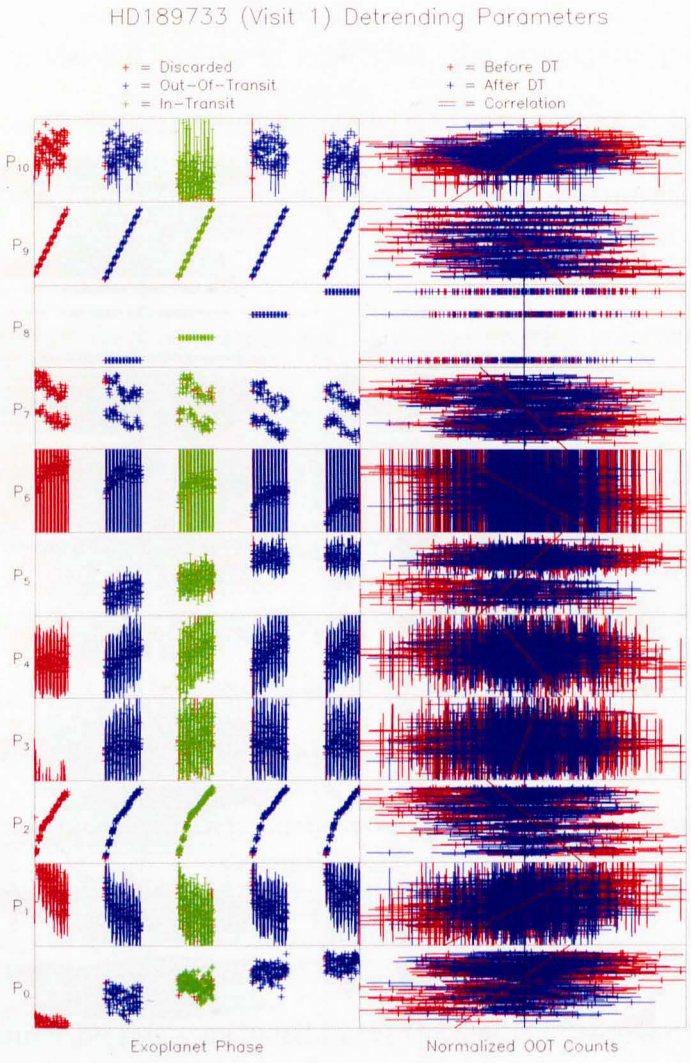


Figure 8.8: *The detrending parameters applied to the secondary eclipse light curve (left) and their modelled correlation with the out-of-transit flux (right). All plots are in the same format as those found in Figure 8.7. Note that the bias parameter (P_{10}) is offset in the orbit containing the secondary eclipse - likely resulting from persistance in the out-of-transit exposures. With such poor coverage of the transit parameter space, the safest option is to omit the bias parameter from the detrending process.*

	Primary Transit	Secondary Eclipse
Before Detrending	5.42×10^{-4}	5.63×10^{-4}
After Detrending	4.22×10^{-4}	4.11×10^{-4}

Table 8.3: *The RMS Spread of the two HD189733 light curves before and after detrending.*

The resulting detrended light curves can be found in Figures 8.9 and 8.10 - shown both in comparison with their raw versions and in isolation. Table 8.3 demonstrates the improvement in the out-of-transit RMS Spread produced by the detrending process; both cases resulting in a significant improvement.

From the comparative plot of the primary transit (Figure 8.9), it is clear that the downward trend noted in Figure 8.5 has been removed; leaving a flatter transit-floor, but preserving curvature as expected. Aside from a visual inspection of the light curve, Figure 8.7 demonstrates negligible residual correlation across all systematic parameters after detrending. The final OOT RMS spread of the primary transit light curve has improved to a value $< 3\sigma_{Poisson}$, approaching the theoretical 'perfect' detrending figure of $2\sigma_{Poisson}$ detailed in Chapter 4. This distribution of points and the precision of this finished light curve allows us to be confident that any significant systematic effects have been removed or sufficiently reduced to allow robust modeling of this transit light curve.

The secondary eclipse (Figure 8.10) can be seen as a clear displacement of the third orbit. However, in this case, the secondary eclipse is not sufficiently distinct for any further modeling to be considered worthwhile. In comparison, Swain et al. (2009b) achieve a superior S/N following the methods detailed in

Swain et al. (2008) (supplementary information). However, attempts to duplicate their analysis proved unsuccessful in acheiving the same S/N ratio[1].

8.5 Light Curve Modeling

The detrended, primary transit light curve can now be modelled to determine an ideal solution for the properties of the transiting system which produced it. The MCMC routine detailed in Chapter 5 was applied with the initial estimates previously listed in Table 8.1 (and which can also be found in Table 8.4).

With these initial conditions, the MCMC began with a correlation length of 73.2 iterations - a noteable improvement on the starting length of the WASP-2 chain (see Chapter 7). This improvement can be attributed to a superior set of intial conditions, since the starting estimates for the chain were based on the results of previous analysis on this same data set (Swain et al., 2008). Once the step size in each dimension of parameter space had been rescaled to produce an overall acceptance rate of 25% and parameter space rotated into a spherically symmetric coordinate system, that correlation length fell to 1.20 iterations (see Figure 8.11).

With a minimised correlation length, the simulation continued until the ideal solution converged. The resulting parameter distributions can be found in Figure 8.12. The most notable relationship in this plot is that between the planetary and stellar radii (R_P and R_S). The degeneracy in this relationship is similar to that identified in the WASP-2 MCMC distributions. However, in this instance the

[1]This may be the result of a lack of comprehensive details supplied in Swain et al. (2008) (supplementary information) or of changes to the HST data processing pipeline made in 2009.

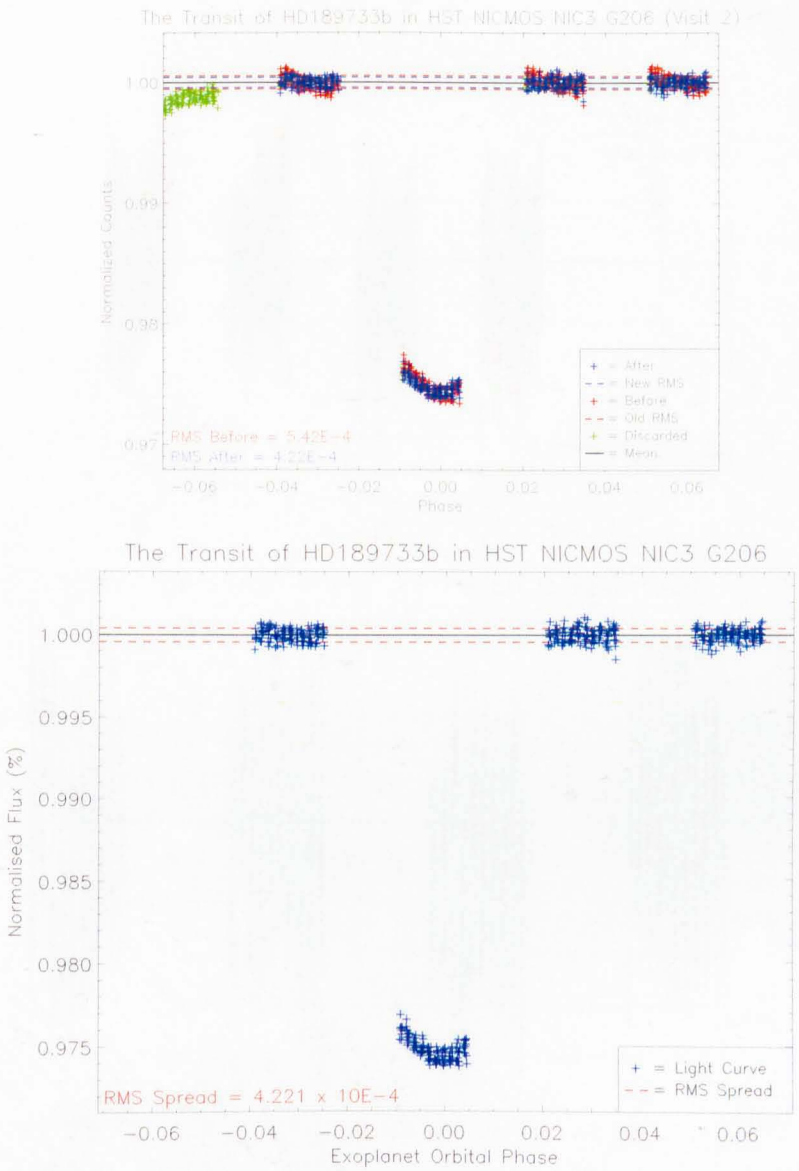


Figure 8.9: *The primary transit light curve after the removal of systematic noise by a process of Preferential Decorrelation (see Chapter 4). The upper plot shows the light curve before (red) and after (blue) detrending, with points removed from processing coloured green. The dashed lines represent the respective RMS Spread of the light curves. The lower plot shows the detrended light curve alone - it is this curve which will be used to model the planetary parameters.*

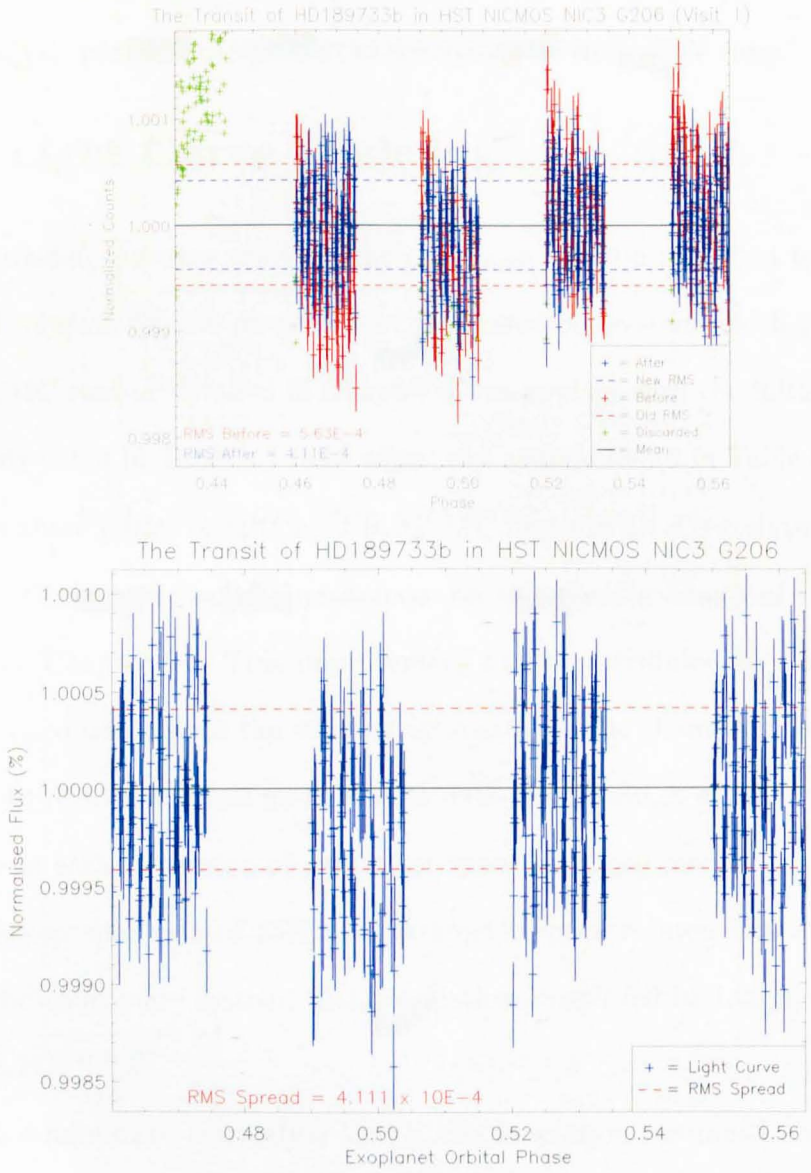


Figure 8.10: *The secondary eclipse light curve with systematic effects removed. Both plots share the same format as found in Figure 8.9. While the secondary eclipse event is evident in the third orbit of observation, it is not distinct enough to allow derivation of a useful model light curve.*

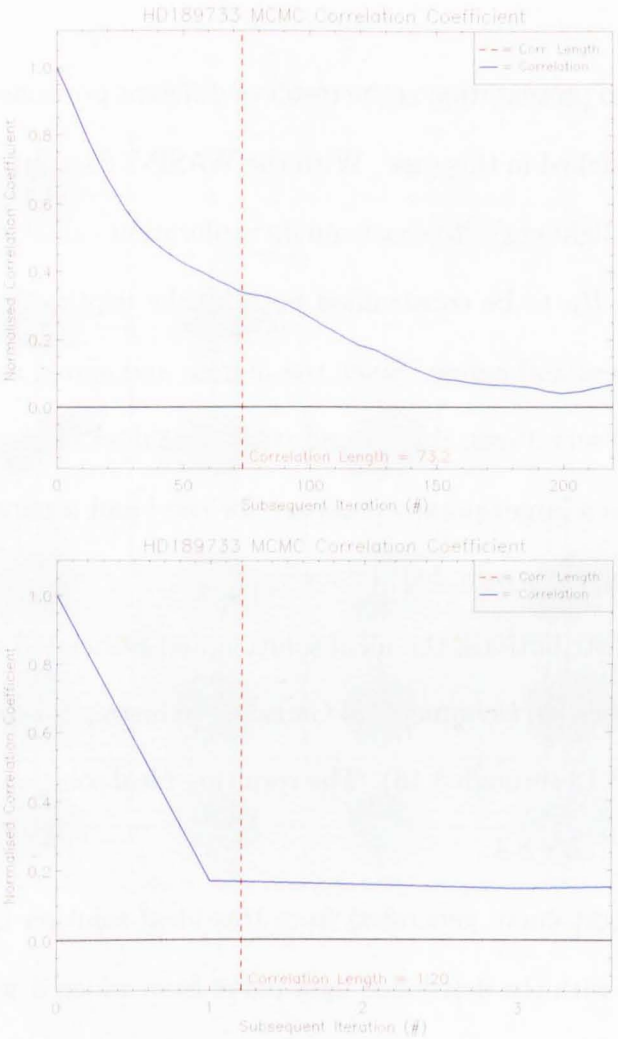


Figure 8.11: *The mean correlation coefficient between any given trial in the MCMC chain and those trials subsequent to it. The upper plot, calculated at the start of the MCMC simulation, shows the evident inheritance between subsequent links in the MCMC chain - demonstrating the inefficiency of the exploration of parameter space. In contrast, the lower plot shows clear improvement once the exoploration has been optimised by the methods described in Chapter 5.*

effect of the degeneracy appears in the radius parameters, rather than between i and a.

This change in presentation is the result of different portions of the incomplete transit being modelled in this case. With the WASP-2 data set, the MCMC chain had a full transit light curve to constrain its exploration - allowing the relationship between R_S and R_P to be constrained both by the depth of transit and by the slope of the ingress and egress. Since the ingress and egress are not captured in the HD189733 observations, this second constraint does not exist. The resulting degeneracy yields a larger parameter space for the i and a parameters to explore - thereby masking their own degeneracy[1].

From these distributions, the ideal solution and associated uncertainties were calculated by fitting an asymmetrical Gaussian to histograms of the accepted trials (see Figures 8.13 through 8.16). The resulting ideal solution and uncertainties can be found in Table 8.4.

The model light curve generated from this ideal solution is shown in Figure 8.17 overplotted with the detrended light curve from which it was discerned. The quality of the fit is perhaps best demonstrated by the associated residuals plot; which demonstrates minimal variation in the third orbit once the model has been subtracted from the detrended light curve.

Comparing these results to those of the WASP-2 data set shows a clear detriment in the uncertainties of both the planetary and stellar radii (R_S and R_P in

[1]Because of its greater inclination angle (i) we would also expect this degeneracy to be greater in WASP-2 than HD189733. However, the two parameters would be expected to demonstrate degeneracy in any such model light curve due to their interdependency (see Chapter 5).

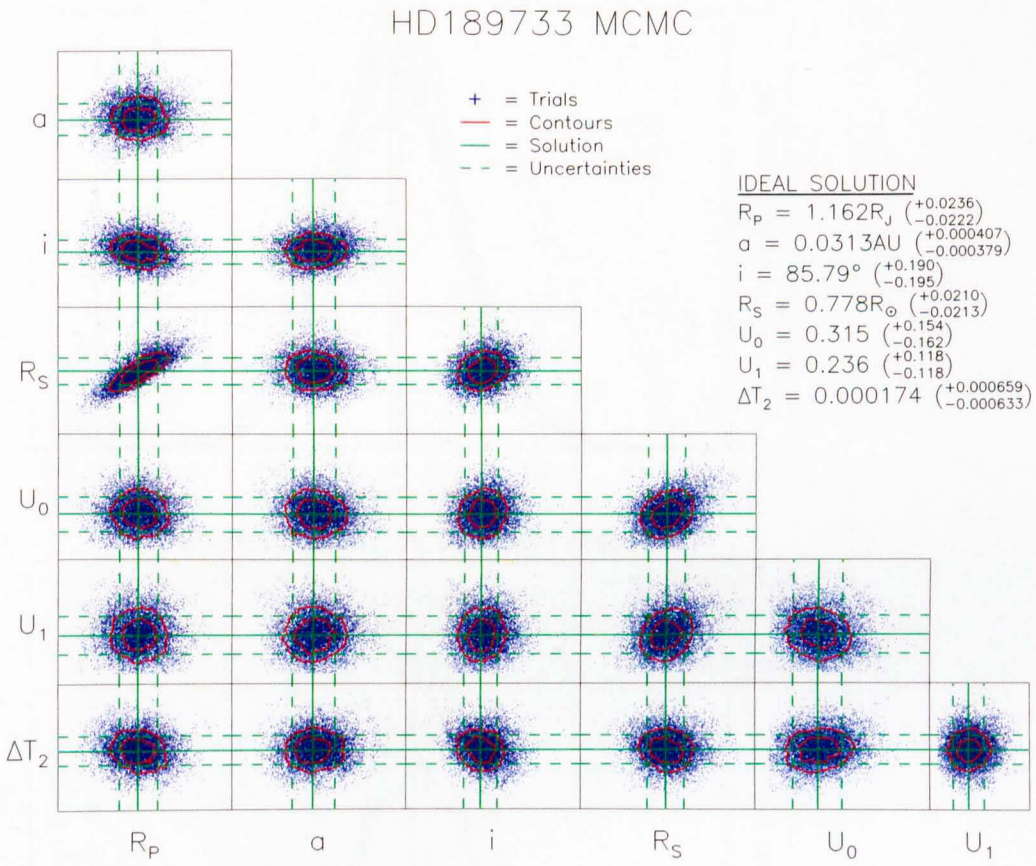


Figure 8.12: *The distribution of accepted trials (blue) across a complete MCMC simulation of the HD189733b transit light curve. The comparative plots demonstrate any dependencies between the parameters as deviations from a uniform circular distribution. In this instance, the Planetary and Stellar Radii (R_P and R_S) exibit an eliptical distribution, betraying a degeneracy between the parameters. The overplotted solid green line represents the value of each parameter in the ideal solution calculated from this MCMC chain and the related dashed lines show the uncertainty on either side of that solution.*

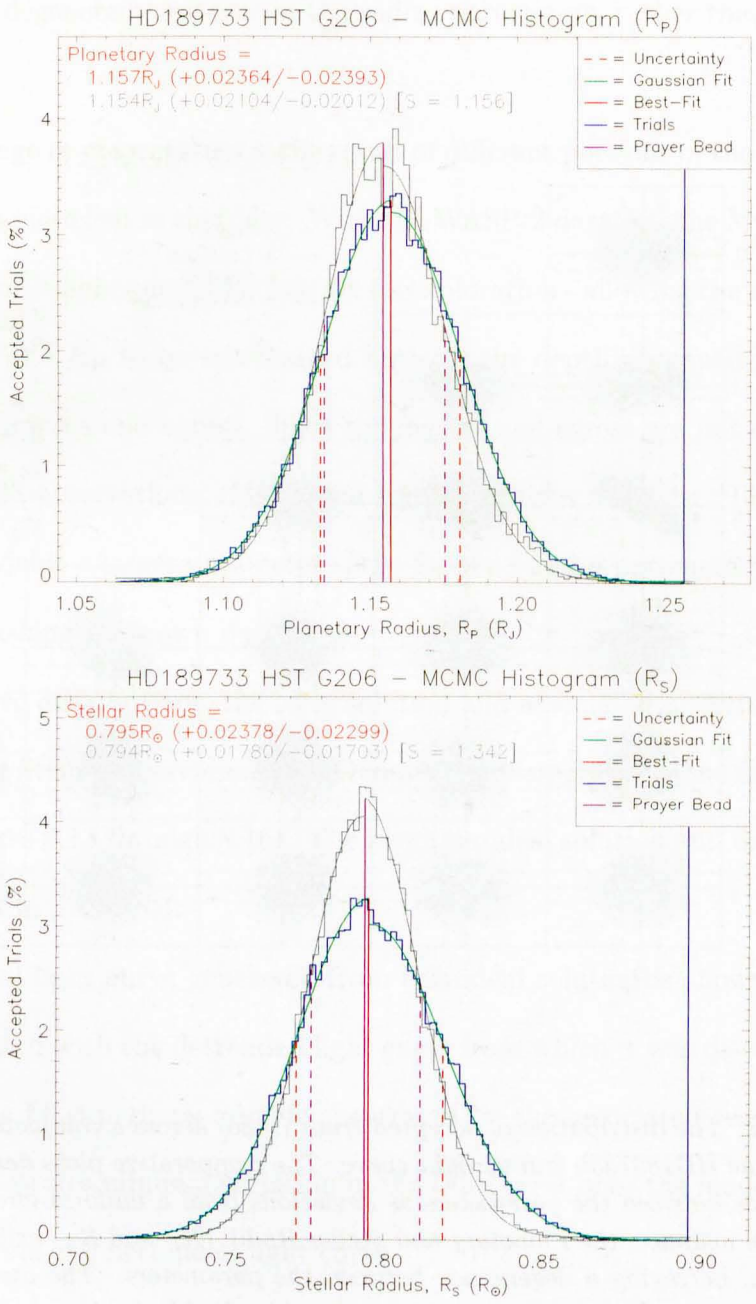


Figure 8.13: *The histograms (blue) of accepted trials in the Planetary and Stellar Radii parameters (R_P and R_S) of a complete MCMC simulation of the HD189733b transit light curve. The overplotted assymetric Gaussian curve (green) models the probability distribution of the MCMC parameter space allowing the ideal solution (red, solid) to be derived along with related uncertainties (red, dashed).*

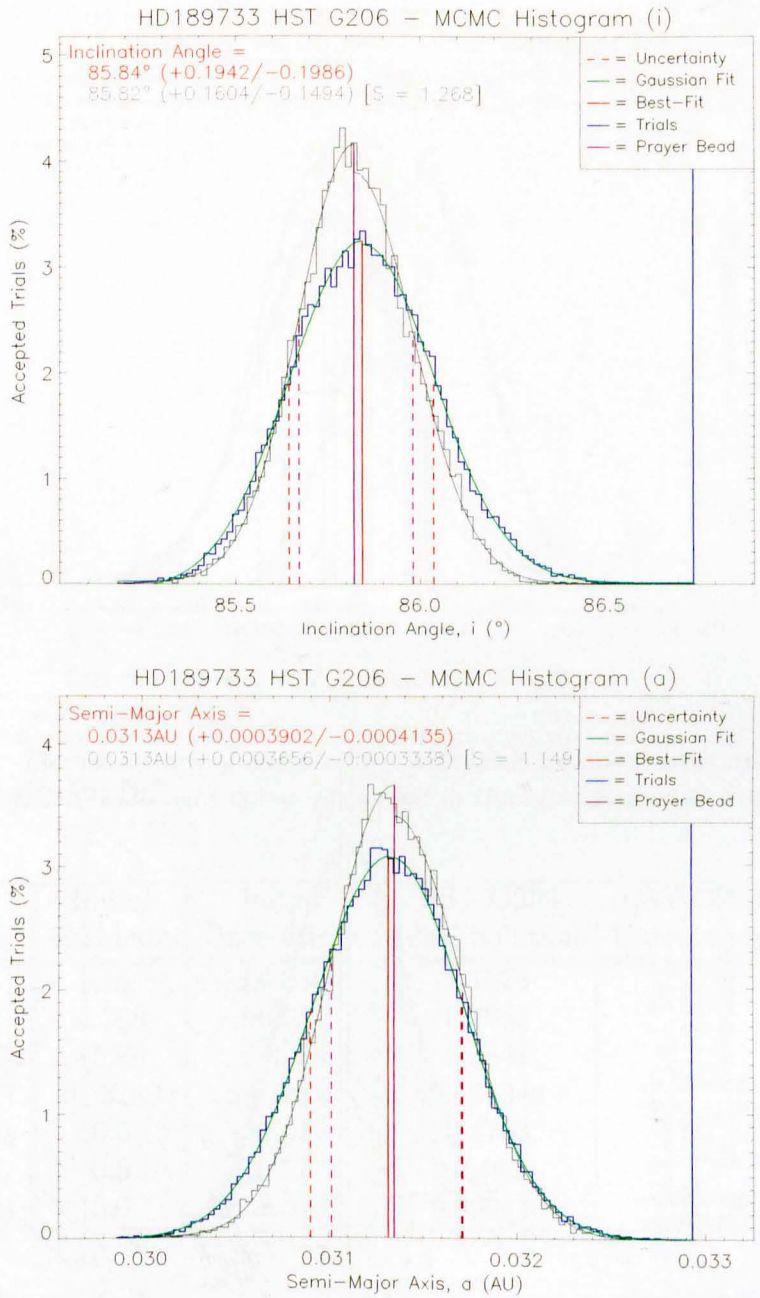


Figure 8.14: *Histograms of accepted MCMC trials in the inclination angle (i) and semi-major axis (a) parameters for the model HD189733b transit light curve, presented in the same format as Figure 8.13.*

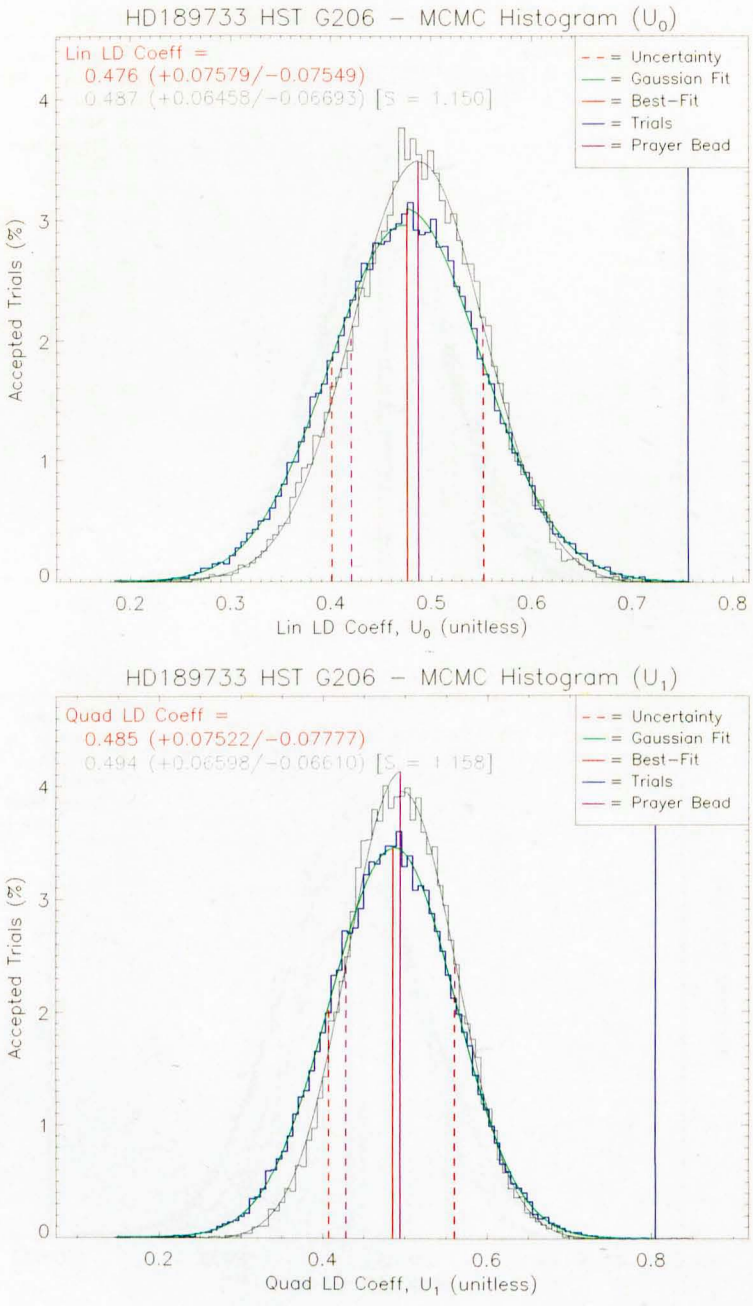


Figure 8.15: *Histograms of accepted MCMC trials in the two limb darkening parameters (U_0 and U_1) for the model HD189733b transit light curve, presented in the same format as Figure 8.13.*

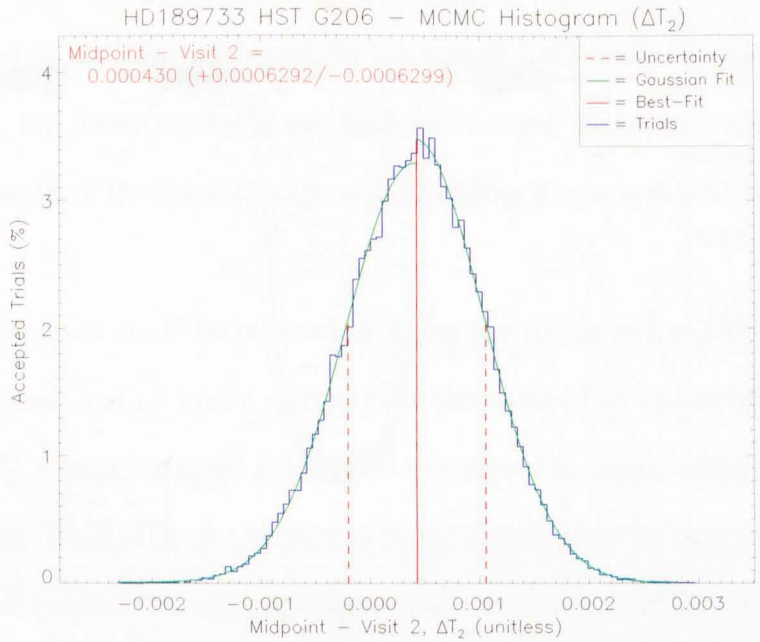


Figure 8.16: *The histogram of accepted MCMC trials for the transit midpoint (ΔT) of the model HD189733b light curve, presented in the same format as Figure 8.13.*

Parameter	Initial Estimate	Initial Uncertainty	HST G206 Ideal Solution	HST G206 Uncertainty	S Value
R_P (R_J)	1.151	±0.038	1.1573	$^{+0.02364}_{-0.02393}$	1.156
R_S $(R_\odot)$	0.788	±0.051	0.7952	$^{+0.02378}_{-0.02299}$	1.342
i (°)	85.76	±0.29	85.841	$^{+0.1942}_{-0.1986}$	1.268
a (AU)	0.03142	±0.0005	0.031342	$^{+0.0003903}_{-0.0004135}$	1.149
U_0	0.5	±0.1	0.4763	$^{+0.07579}_{-0.07549}$	1.150
U_1	0.5	±0.1	0.485	$^{+0.07522}_{-0.07777}$	1.158
ΔT $(\Delta\phi)$	0.0	±0.00001	0.000430	$^{+0.0006292}_{-0.0006299}$	N/A

Table 8.4: *The ideal solution of the MCMC simulation of the HD189733b transit light curve shown in contrast to the initial estimates supplied to that process. With no reliable figures attributable to the limb darkening parameters, the values of U_0 and U_1 were started with initial estimates of 0.5. This resulted in a longer than normal MCMC chain while the parameters located around the ideal solution and rescaled their step-size accordingly, however, the final results appear sound.*

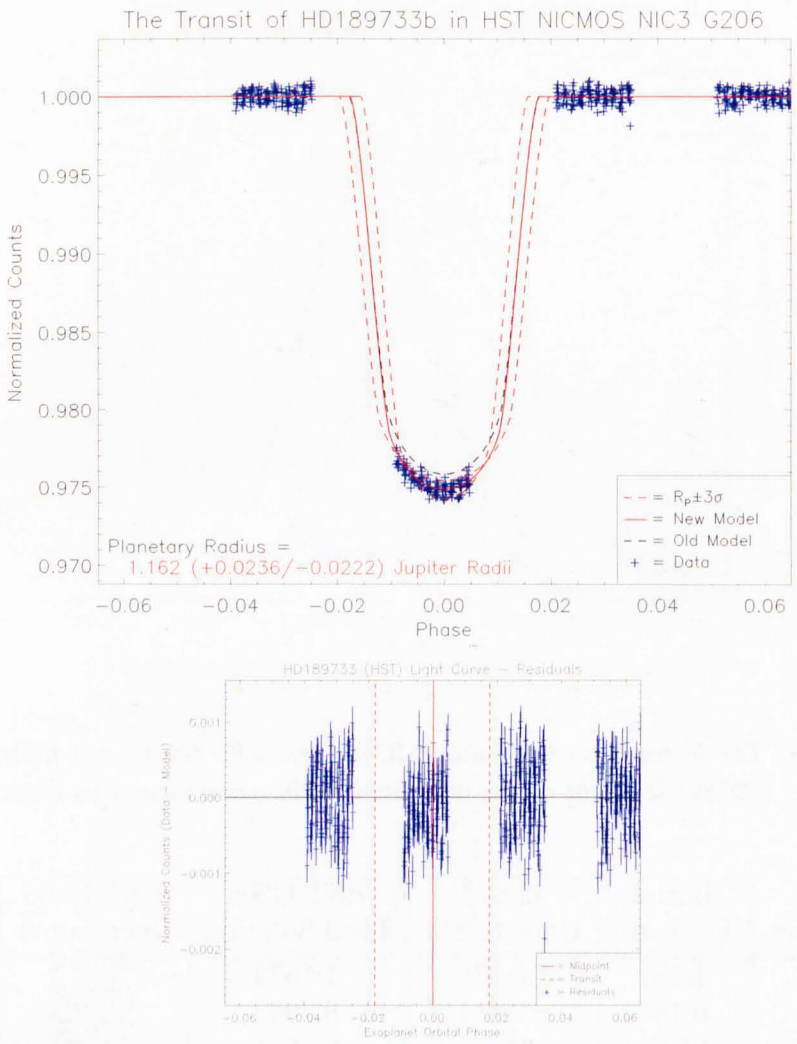


Figure 8.17: *The model light curve for the transit of HD189733b as observed by HST NICMOS. The upper plot shows the measured, detrended light curve (blue) overplotted with the model created from the ideal solution of the MCMC simulation (red, solid) and the upper and lower boundaries on that model (red, dashed) calculated from the related uncertainties. The dashed black line shows the light curve created from the initial estimates provided to the MCMC process. The lower plot shows the residuals (blue) once this model is subtracted from the data, with the transit midpoint denoted by the solid red line and the dashed red lines marking out the full duration of the transit event.*

Table 8.4). This is the result of the degeneracy noted in Figure 8.12 and results from the lack of an ingress or egress in the transit light curve. As can be seen in Figure 8.17, the uncertainties of the light curve allow for a large range of variation on the walls of the transit, while still providing a reasonable fit to the transit floor.

This degeneracy could be reduced by fixing the stellar radius parameter (R_S), either at the value of its initial estimate, or the value of an initial MCMC chain. Alternatively, a more stringent prior could be applied to one or both radii parameters. However, both of these approaches would simply serve to restrict the motion of the MCMC chain through parameter space and would artificially reduce the uncertainties by preventing exploration of real solutions to the light curve. The most robust approach to reduce the effect of the degeneracy would be to concurrently model a second HD189733b transit lightcurve covering those portions of the transit not captured by NICMOS. Such a light curve would need to be taken at a similar wavelength to that of the G206 grism to give a comparative value of R_S.

The other parameters have uncertainties comparable to those seen in the WASP-2 data set - as would be expected for light curves with such similar RMS spreads. While the simulation does report a small variation in the transit midpoint (ΔT), the figure is easily consistent with zero and not sufficient to refine the planet's period or ephemeris; nor to determine the presence of any transit timing variations.

Most encouragingly the limb darkening parameters arrived at a good fit to the

light curve (see Figure 8.17 and the histograms in Figure 8.15), despite very poor initial estimates. This helps confirm the reliability of the MCMC algorithm, as well as potentially demonstrating a reliable[1] method of determining stellar limb darkening alongside the transit parameters where previous studies have used fixed values of limb darkening which needed to be estimated in advance.

The prayer beading results (shown in grey and purple in Figures 8.13 through 8.15) reveal a significant degree of systematic noise persisting in-transit. The average S factor for the light curve is 1.204, with the Stellar Radius and Inclination Angle parameters most significantly affected - as was the case with WASP-2. This indicates that 20% of the measured uncertainty in the parameters is the result of residual systematic noise in-transit. However, as the histograms demonstrate, the variation between the Ideal and Prayer Bead Solution is well within the uncertainties and the prayer bead histograms exhibit no significant non-Gaussian structure which would also indicate residual systematic effects.

8.6 Atmospheric Analysis

As reported in Table 8.2, the width of the PSF in the observations of the HD189733 primary transit is $9.63 pixels$ (or $0.111 \mu m$ in the spectral domain). Using this as our minimum bin-width, a series of overlapping bins were constructed, spanning the full length of the light gathering box (again, see Table 8.2), resulting in 53 wavelength-dependent transit light curves from 1.645 to $2.242 \mu m$.

[1] As was shown with the WASP-2 data set (see Chapter 7), the MCMC modelled limb darkening values closely matched the initial estimates calculated by Charbonneau et al. (2007).

These wavelength-dependent curves were then modelled alongside the overall light curve to determine any variation in transit depth (i.e. variation in the apparent radius of the planet, R_P) with wavelength. A sample HD189733 wavelength-dependent transit light curve can be found in Figure 8.18, overplotted with the model light curve resulting from this process.

The reliability of the wavelength-dependent light curves is demonstrated in the comparative out-of-transit RMS spread of the curves across the first order spectrum. Figure 8.19 reveals one residual systematic feature which is not entirely removed by the detrending process. As a result, any apparent features in transmission spectrum which fall in the 1.89-1.96μm range can not be believed.

The complete set of wavelength-dependent model light curves are perhaps best presented in Figure 8.20 as a continuum plot. The exoplanet's atmospheric transmission spectrum can then be extracted by converting the variation in the planetary radius (R_P) into the relative absorption of the planet ($(R_P/R_S)^2$). The resulting spectrum can be found in Figure 8.21; shown both in isolation and alongside the previously published results by Swain et al. (2008) and Sing et al. (2009).

8.6.1 Residual Systematics

The variation in transit depth with wavelength shows several clear features, but before any analysis can take place, we must satisfy ourselves that these features are not the result of residual systematic effects. Firstly, the average depth of the wavelength-dependent light curves matches that of the overall light curve (shown

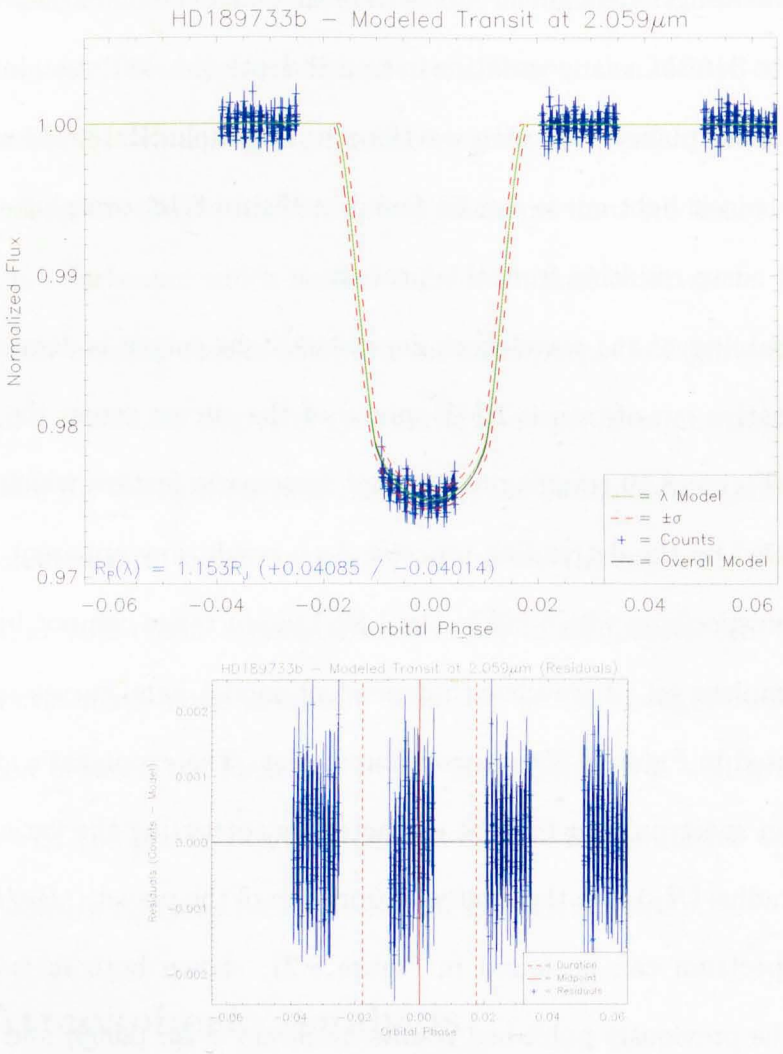


Figure 8.18: *A wavelength-dependent HD189733 transit light curve (top, blue) taken from the centre of the first order spectrum at 2.059μm. After detrending the RMS spread of the out-of-transit points was measured at 5.99×10^{-4}. The green line demonstrates the ideal solution of the MCMC process used to fit the curve, with a transit depth slightly less than that of the overall light curve (whose model is demonstrated by the solid black line). The dashed red lines show the 1σ uncertainties on the value of $\Delta R_P(\lambda)$ combined with the uncertainty on R_P measured from the model of the overall light curve. The residuals of the fit (bottom, blue) show no residual features in the transit orbit - demonstrating a good solution. The solid red line marks the transit midpoint and the dashed red lines the transit duration.*

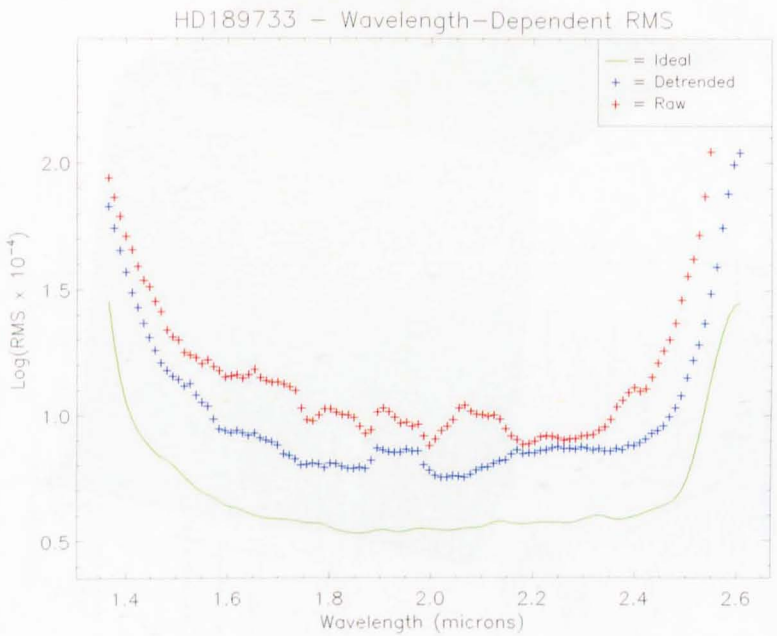


Figure 8.19: *The variation in out-of-transit RMS spread of the wavelength-dependent transit light curves across the first order spectrum both before (red) and after (blue) detrending. The green line shows the theoretical limit of a 'perfect' detrending process (see Chapter 4 for a full discussion). While the majority of the features in the red curve are removed during detrending, one large feature remains - likely the result of either an unknown systematic effect or a poorly constrained systematic.*

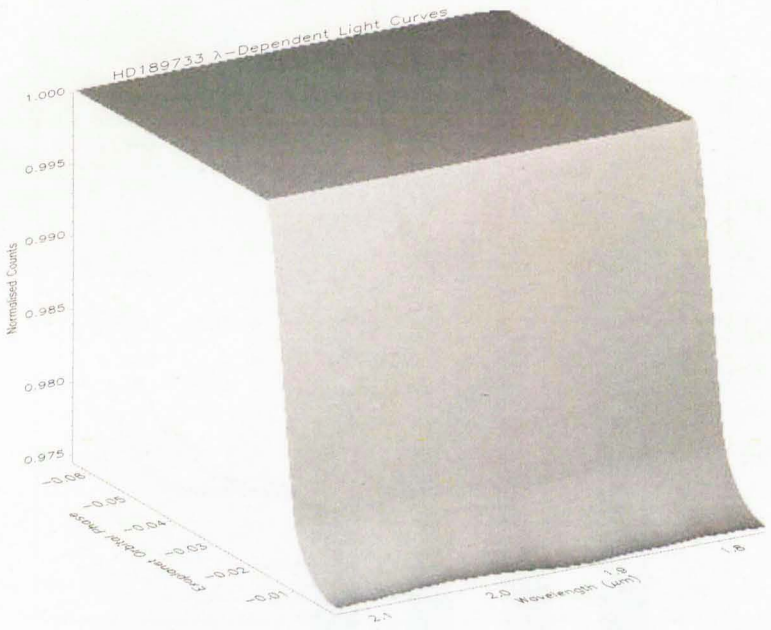


Figure 8.20: *A continuum plot of the models generated for every wavelength-dependent light curve across the first order spectrum of the HD189733 data set. In generating these models, those parameters with no expected wavelength-dependence are constrained by modeling the overall light curve. The wavelength dependent parameters are modelled independently for each wavelength-dependent light curve (though this process is performed in parallel with the modeling of the overal light curve, see Chapter 6). Small variations in transit depth are noticable at the base of the transit continuum. See Figure 8.18 for details on the uncertainty of these models.*

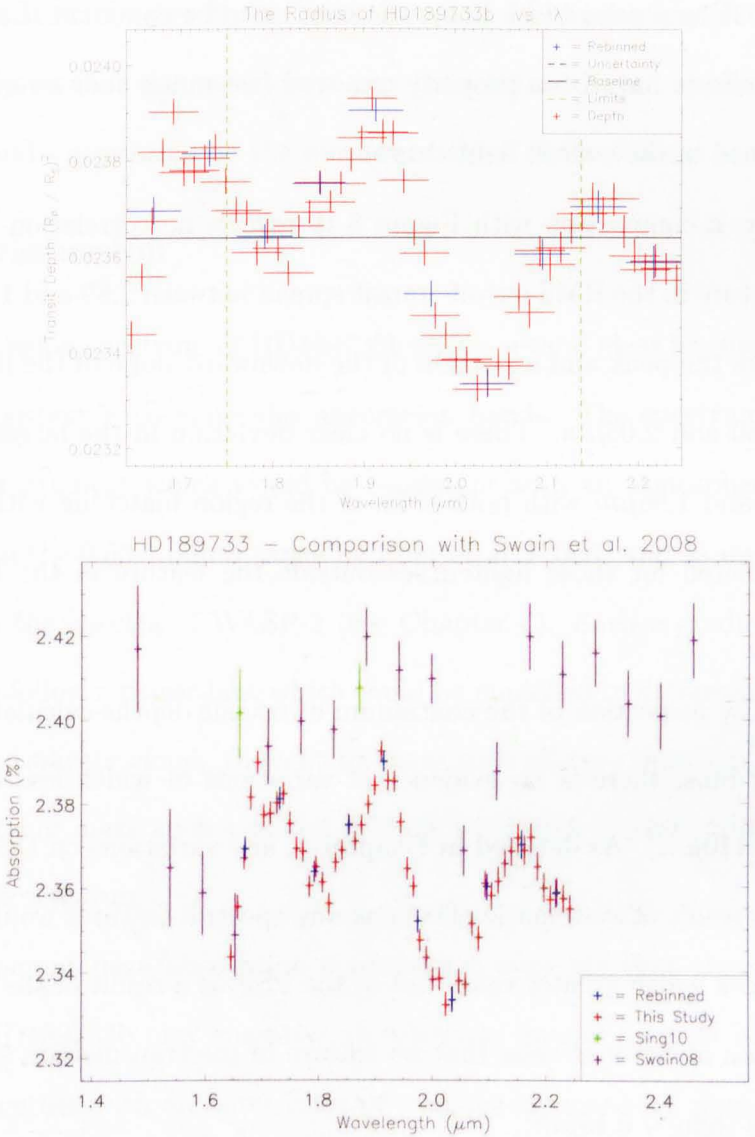


Figure 8.21: *The variation in HD189733b's transit depth with wavelength. The upper plot shows the results of this study in isolation, while the lower plot shows them in conjunction with the results of Swain et al. (2008) and Sing et al. (2009). The upper plot also shows the depth of transit observed in the overall light curve (black, solid; see Figure 8.17) and the uncertainty on this value (black, dashed). Both plots show the measured* $(R_P/R_S)^2$ *for each wavelength-dependent light curve (red) and the average of those depths re-binned to the width of the PSF (blue). The transmission spectrum shows several clear features, none of which match with the residual features in the RMS of the modelled light curves (see Figure 8.19).*

in Figure 8.21 as a solid black line), allowing us to be confident that any global systematic effects have been properly removed (assuming that no systematic effects remained in the overall light curve).

Secondly, a comparison with Figure 8.19 reveals no correlation with the remaining feature in the RMS out-of-transit spread between 1.89 and 1.96μm. This region covers the peak and a portion of the downward slope of the larger feature between 1.80 and 2.05μm. There is no clear deviation in the larger feature between 1.89 and 1.96μm with both sides of the region matching with the transit depth calculated for those lightcurves outside the feature in the RMS spread curve.

Finally, by inspection of the continuum of transit depths calculated from the overlapping bins, there is no evidence of variations of width less than that of the PSF (0.110μm). As detailed in Chapter 6, any variations on this scale could only be the result of systematic effects as any spectral features would have been smoothed to a width greater than that of the PSF as a result of the defocussing. By inspection alone, it is clear that no feature in the transmission spectrum has a width less than $\sim 0.15\mu m$[1].

It should be noted that systematic effects are visible in the results from individual light curves. Performing a more detailed comparison with the RMS spread distribution (see Figure 8.19), the light curve centred at $\sim 2.16\mu m$ has

[1]In regard to systematic effects, we must consider any specific feature as either a step-change or a peak/trough in its entirety. Evidence of systematics in this regard would be a region of points offset by some arbitrary value while continuing the same trend as the points on either side; or a peak or trough in the data, again with the unaffected points on either side of the region continuing the same trend.

a detrended RMS greater than those around it. Examination of the resulting spectrum shows that this same curve reports a transit depth apparently greater than it should - appearing on the downward slope of the feature at 2.15.

8.6.2 Discussion

The transmission spectrum of HD189733b shows several clear features with magnitudes consistent with molecular absorption bands. The spectrum also shows an apparent gradient which would be consistent with an atmospheric haze - as was found in the 0.55-1.05μm range by Pont et al. (2008) and as was tentatively observed in the spectra of WASP-2 (see Chapter 7). Such a gradient would be expected to follow a power-law, which could be modelled to determine the sizes of grains in the silicate clouds thought to cause such effects. However, the features in the spectrum make such a model difficult without a corresponding model for the absorption features.

The apparent haze/absorption combination suggests that the upper atmosphere of HD189733b may comprise an relatively opaque layer of haze and condensates together with an outer layer of transparent material, potentially being boiled off the planet by incident radiation from the star - though producing a smaller variation in transit depth to that found in the WASP-2 data set, despite the two planets having comparable Roche lobes (see Figure 8.22). Such a scenario would be expected to produce a slope, as found by Pont et al. (2008) and as appears in Figure 8.21, with features imprinted on top of that slope by absorption in the hypothetical transparent, upper layer. Spectroscopic modeling is required

to draw further, more comprehensive conclusions.

The feature in Figure 8.21 between 1.6 and 2.0μm would be consistent with water absorption were it not for the dip just below 1.8μm. However, if the spectral features are produced by material being boiled off the atmosphere, we would expect to see emission features which would be consistent with this feature. Alternatively, the observed feature could be a combination of water absorption and that of other molecules at 1.7 and 1.9 microns.

In comparison with the results of Swain et al. (2008) and Sing et al. (2009), the apparent difference in $(R_P/R_S)^2$ between the Sing et al. (2009) results and our own is most likely the result of different limb-darkening models. Sing et al. (2009) apply a more complex model to address the variation of intensity across the disk of the star, while the limb darkening in our own model light curve is governed by a simple, quadratic law (see Chapter 5).

While we can be satisfied that the limb darkening coefficients derived by this research are robust[1], it is possible to change the measured $(R_P/R_S)^2$ by $\sim 0.1\%$ by reducing each limb darkening coefficient by $\sim 2\sigma$. While this would not impact the features observed in the spectrum (the variation of limb darkening with wavelength is expected to be small in the infrared) it makes the average $(R_P/R_S)^2$ in any set of results difficult to compare with other studies at this level of precision. Such a comparison could be made if the wavelength-dependent modeling was performed with the limb-darkening coefficients fixed at the same

[1]The MCMC light curve simulation software allows both limb darkening coefficients U_0 and U_1 to vary - meaning each arrives at the value which provides the most probable fit to the data. See Chapter 5 for a full explanation.

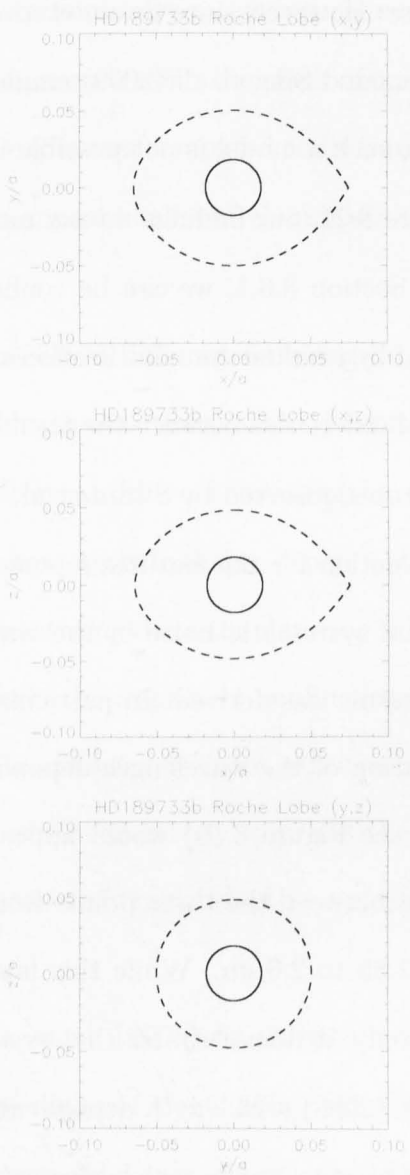


Figure 8.22: *The disk of HD189733b (solid) shown in comparison with the cross section of its Roche lobe (dashed). The upper plot shows the cross-section in the orbital plane of the system. The middle plot shows the cross-section in the plane perpendicular to the orbital plane and containing our line-of-sight at mid-transit (in both cases, the parent star is to the right of the plot). The lower plot shows the cross section along our line-of-sight at mid-transit - i.e. the cross-section observed during transit. (Busuttil & Haswell, 2011, in prep; and poster presented at NAM2011.)*

values employed elsewhere. However, since Swain et al. (2008) do not state their limb darkening parameters and Sing et al. (2009) employ a different model with incompatible parameters, such a model is not possible in this case.

As is clear from Figure 8.21, our findings do not match those of Swain et al. (2008). As discussed in Section 8.6.1, we can be confident that our results are not significantly impacted by residual systematic effects. Therefore, to determine the source of the observed differences between the results, we must consider what may have caused the features observed by Swain et al. (2008).

The most likely explanation for the features reported by Swain et al. (2008) is the presence of residual systematic noise in the wavelength-dependent light curves from which their result was derived. In particular, the residual feature in the RMS spread distribution of the wavelength-dependent light curves (though generated in this study, see Figure 8.19) would appear to match with a step-change in the Swain data between the three points from 1.7 to 1.85μm and the subsequent points from 1.85 to 2.0μm. While this feature is not significant to the conclusions of that study, it demonstrates that systematic noise may indeed persist in the Swain et al. (2008) wavelength dependent light curves.

However, another explanation exists which also serves to explain the shape of the Swain et al. (2008) transmission spectrum. The NICMOS instrument is known to suffer from nonlinearity - an effect whereby any photon falling onto the detector is less likely to be counted if the pixel in question is already carrying a significant charge. As a result, pixels exposed to a large amount of flux will report ever-so-slightly fewer photons than they otherwise would, since photons arriving

towards the end of the exposure have a reduced chance of being detected.

In the case of an explanetary transit, the effect of nonlinearity would be greater in out-of-transit (OOT) exposures compared to in-transit exposures. This would cause the transit itself to appear shallower than it should - though only by a very small amount.

When considering transits at different wavelengths, another factor comes into play - namely, the wavelength sensitivity of the detector. Since the NIC3 camera more readily detects photons at different wavelengths, the first order spectrum in each grism image has a characteristic shape, dependent on the applied grism. For parts of the first order spectrum where the detector is less sensitive, less flux will be recorded and, therefore, the impact of nonlinearity is diminished. Similarly, at wavelengths where the detector is more sensitive, the effect of nonlinearity is accentuated. This contrast effectively imprints the wavelength sensitivity of the detector onto the measured transit-depth - unless the detector nonlinearity is properly accounted for.

Comparing the Swain et al. (2008) result with the wavelength sensitivity of the NIC3 detector in the range of the G206 grism, there is a clear similarity in the shape of the two curves. The only significant difference occurs in the points around $2.1\mu m$ which, in the Swain et al. (2008) model spectrum, bridge the gap between the water and methane absorption features.

Looking more closely at the detector, however, it should be noted that these points are taken from pixels lying on the edge of one of NIC3's four CCD chips in a region where the dark current is known to be both erratic and temperature-

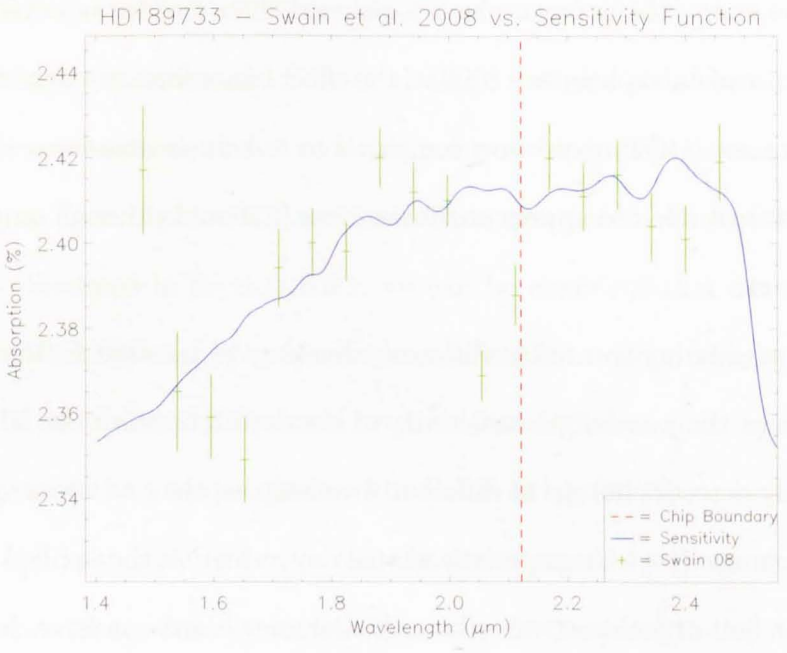


Figure 8.23: *The HD189733b atmospheric transmission spectrum derived by Swain et al. (2008) (green) in contrast with the inverse sensitivity function of the NIC3 detector (blue) which has been scaled to the data for comparison. The dashed red line shows the edge of one of the four CCD chips in the NIC3 detector. Several columns ($\sim$ 0.1-0.2μm) to the right of that boundary (to the left in wavelength-space) are known to be prone to variations in dark current.*

dependent (Skinner et al., 1998). As such, these points could easily have been differently impacted by the nonlinearity effect, as well as likely being prone to exaggerated systematic effects. Figure 8.23 shows the Swain et al. (2008) result in conjunction with a scaled plot of the detector sensitivity.

Prior to 2009, the automated HST data pipeline (which processes the raw images prior to their delivery to the researcher who requested them, see Chapter 2) had a different in-built mechanisms to account for detector nonlinearity, temperature dependence and other potential sources of systemetics. New routines were

introduced in 2009, potentially explaining why this study failed to reproduce the 2008 result.

Given the clear similarities evident in Figure 8.23 and the failure of both Gibson et al. (2011) and this study to duplicate the result, along with the null detections reported by Sing et al. (2009) and Mandell et al. (2011), both searching for Methane with different instruments, it seems highly likely that the detections of Swain et al. (2008) are erroneous.

8.7 Conclusions

The HD189733 data set has produced an interesting and believable transmission spectrum (see Figure 8.21) which can now be analysed by atmospheric modellers. The spectrum shows no evidence of residual systematics when compared to both the distribution of post-detrending RMS spreads with wavelength (see Figure 8.19) and when checked for step-changes in the overlapping wavelength-dependent bins.

These results have also yielded important lessons for future research. Firstly, the lack of a transit ingress or egress was seen to create a large degeneracy between the stellar and planetary radii (R_S and R_P) during the detrending process (see Section 8.5). The planetary parameters derived from the MCMC process would likely have been far better constrained had a second visit been used to capture those sections of the transit not captured in the first. Furthermore, checking that the transits in two such visits overlap when phase-folded would provide a further check on the veracity of the detrending process.

Finally, analysis of this data has proved the importance of a reliable and robust detrending process in relation to the Swain et al. (2008) result. As previously demonstrated by Gibson et al. (2011), the transmission spectrum of Swain et al. (2008) is greatly dependent on the method of systematic noise reduction. In contrast, the results of this study have two checks on the detrending process (as detailed in Section 8.6.1). Furthermore, the same robust pipeline has been applied to multiple data sets (see Chapters 7 and 9) providing a more comprehensive validation of the methods involved.

Chapter 9

Results - GJ436

9.1 Chapter Outline

The following chapter details the results of the GJ436b transit data set as derived from the data processing pipeline covered in Chapters 3-6. The chapter presents the current knowledge of the GJ436 system and previous research on the exoplanet and its atmosphere. Following this, pipeline results of the raw, processed and modelled light curves are presented, along with details and anlysis of the processing steps there involved. Finally, the chapter details the wavelenght dependence of the transit depth and discusses the resulting transmission spectrum of the exoplanet's atmosphere.

9.2 Introduction

GJ436 (also called Gliese-436) was one of the first systems found to hold a 'Hot Neptune' exoplanet (i.e. a planet whose orbit has a very small semi-major axis and with a radius and mass comparable to that of Neptune). It was discovered by Butler et al. (2004) via measurement of radial velocity variations in GJ436.

Parameter	Value	Uncertainty
Planetary Radius, R_P (R_J)	0.365	± 0.02
Planetary Mass, M_P (M_J)	0.0737	± 0.0052
Inclination Angle, i (°)	85.8	$\pm^{0.21}_{0.25}$
Semi-Major Axis, a (AU)	0.02887	$\pm 9.5 \times 10^{-4}$
Period, P ($days$)	2.6438986	$\pm 1.6 \times 10^{-6}$
Transit Epoch, E (HJD)	2454222.61588	± 0.00012
Stellar Radius, R_S ($R_\odot$)	0.464	$\pm^{0.009}_{0.011}$

Table 9.1: *Latest planetary and stellar parameters for the GJ436 system, collated from Southworth (2010), Bean et al. (2008), Cáceres et al. (2009) and Torres (2007).*

GJ436b was found to transit its parent star by Gillon et al. (2007b) and quickly became the focus of many follow-up observations to properly constrain the planet's parameters. These observations included the detection of the planet's secondary eclipse - recorded using the Spitzer space telescope by Deming et al. (2007). The most up-to-date planetary parameters were calculated by Southworth (2010) by aggregation of multiple separate observations and can be found in Table 9.1.

As would be expected, GJ436b's far smaller size means it occupies a much different region of the transiting exoplanet population to WASP-2b and HD189733b. As demonstrated in Figure 9.1, GJ436 is only one of a handful of known transiting planets with masses and radii substantially less than that of Jupiter. It is also among the coolest known transiting exoplanets with a temperature estimated to be just above 750K.

GJ436 has long been considered a prime candidate for transit-timing searches (see Chapter 1) which would be more easily discernable than in hot Jupiter sys-

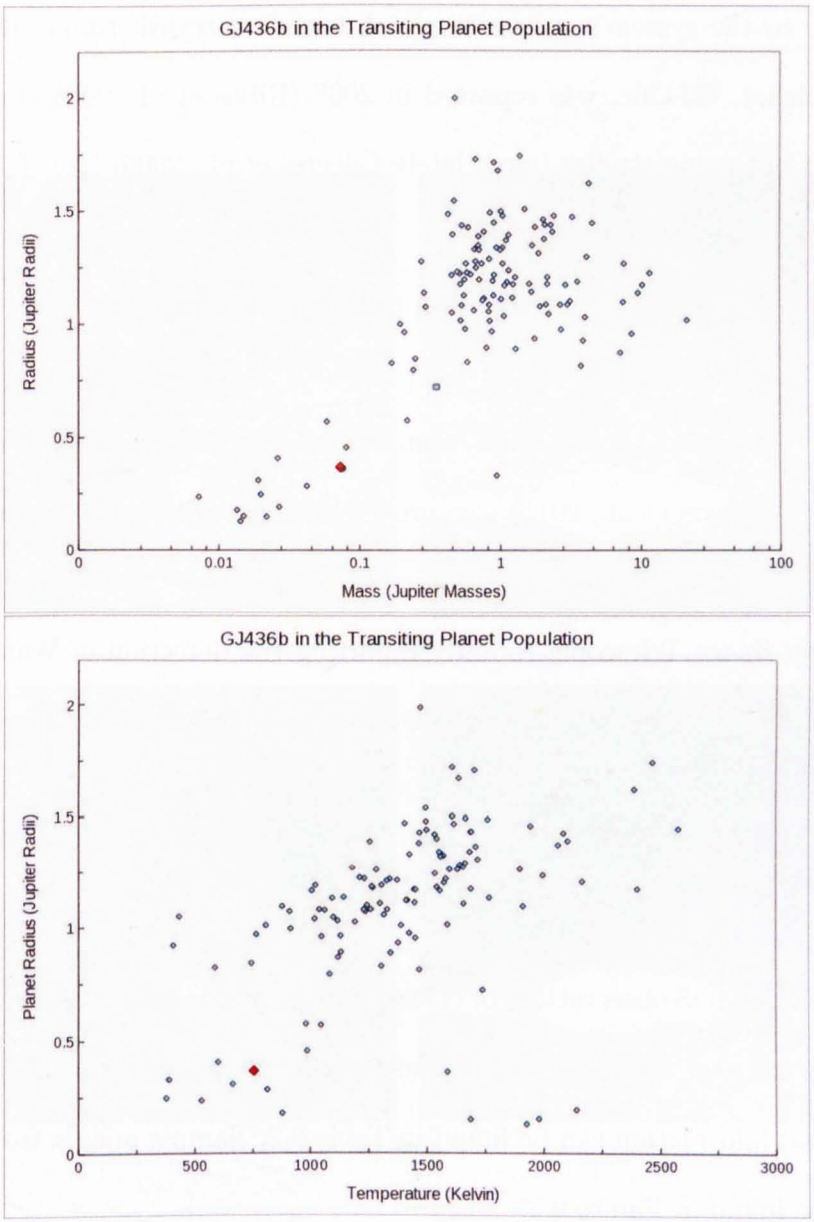


Figure 9.1: *The exoplanet GJ436b (red) in contrast to the population of transiting planets as a whole (blue). The upper plot shows the relation between mass and radius of the planets while the lower plot shows the effect of temperature on the planets' radii. Since the majority of transiting exoplanets discovered to-date are Hot Jupiters, the Hot Neptune GJ436 falls in region far below the densest region of the population.*

tems due to the system's parameters[1]. A tentative transit-timing detection of a third planet, GJ436c, was reported in 2008 (Ribas et al., 2008) but later retracted. Continuing studies (most lately Cáceres et al., 2009) have failed to find any further evidence of GJ436c.

The observations to be reported in this chapter were performed in late 2007 and preliminarily reported in 2009 (Pont et al., 2009) with updated planetary parameters but no detection of any atmospheric features. A further study of the same data (Gibson et al., 2011) also proved inconclusive in this regard.

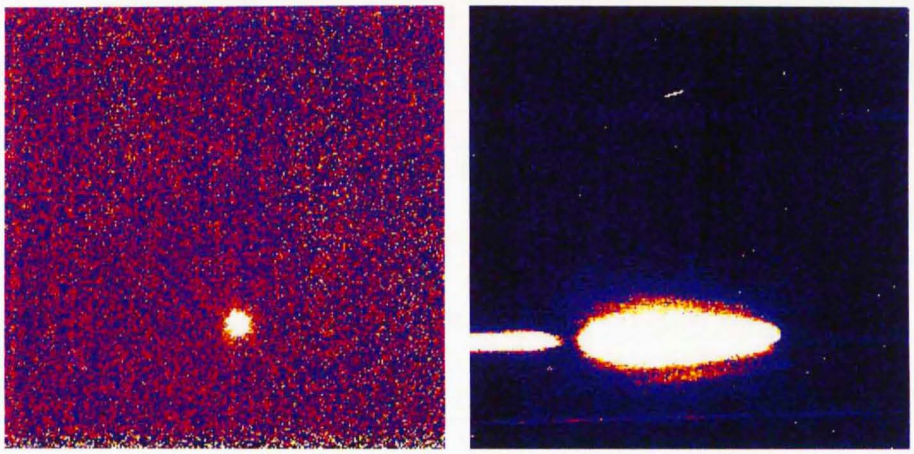
More successful atmospheric studies of GJ436 have been performed using the Spitzer Space Telescope; recently reporting the detection of Water, Carbon Monoxide and Carbon Dioxide (Madhusudhan and Seager, 2011; Stevenson et al., 2010).

9.3 The GJ436 Data Set

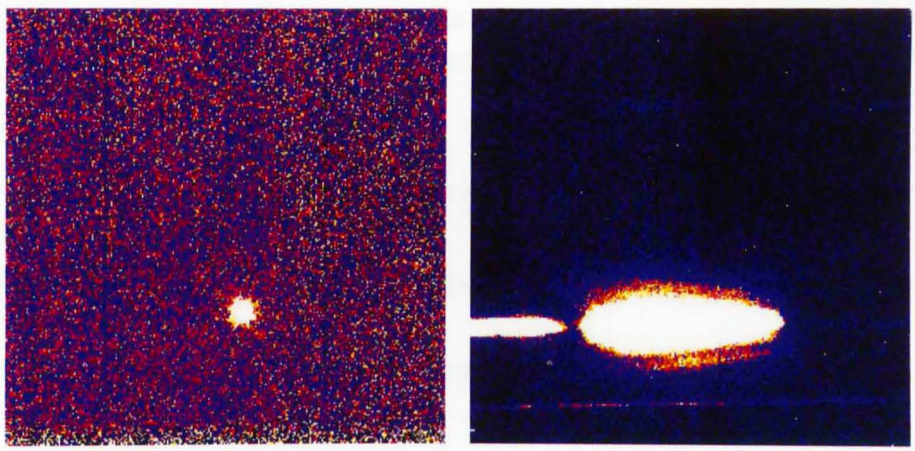
The HST NICMOS observations of GJ436 comprise two visits of observation, each capturing the floor and egress of the transit event. Full details on the observations and related information can be found in Table 9.2. Sample images from the data set can be found in Figure 9.2.

The grism images were flat-fielded by application of a composite flat, derived as detailed in Chapter 3 from a wavelength key determined by location of the

[1]When compared to hot Jupiters, GJ436b has a comparatively smaller mass and is therefore more easily perturbed by other planets in the system. While this would normally be offset by the reduced transit depth of the system leading to an innevitably inferior S/N, GJ436a's radius of only $0.464 M_{\odot}$, combined with an apparent magnitude of 10.68, lessens this detriment considerably.



(a) Visit 1



(b) Visit 2

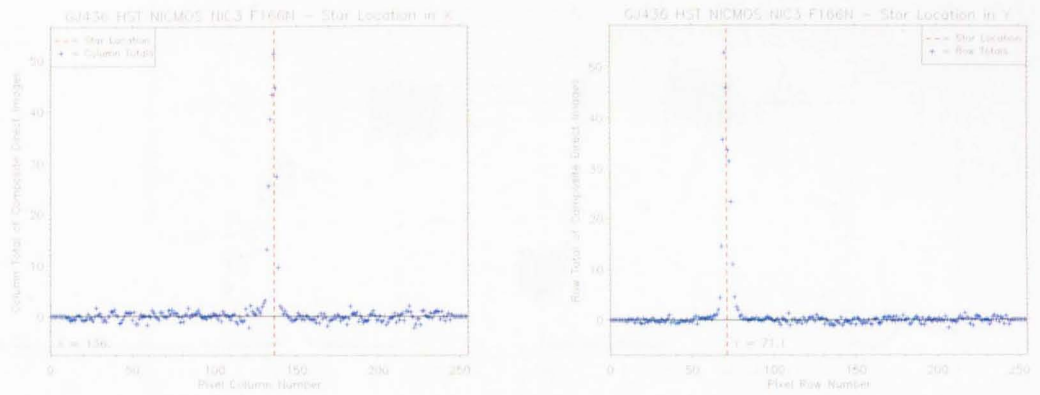
Figure 9.2: *Sample filter (left) and grism (right) exposures from the GJ436 HST NICMOS data set.*

	Visit 1	Visit 2
Containing	Transit Floor & Egress	Transit Floor & Egress
Date of Observations	November 2007	December 2007
Number of Orbits	4	4
Filter	F166N	F166N
Filter Exopsure Time (s)	0.993	0.993
Number of Filter Images	1	1
Grism	G141	G141
Grism Exposure Time (s)	1.990624	1.990624
Number of Grism Images	935	917
Star X Pos ($pixels$)	136.24	138.01
Star Y Pos ($pixels$)	71.16	76.08
Box X Dimensions ($pixels$)	60.40 : 185.26	62.18 : 187.04
Box Y Dimensions ($pixels$)	51.66 : 76.18	57.12 : 80.77
PSF Width ($pixels/\mu m$)	9.84 / 0.079	9.41 / 0.075

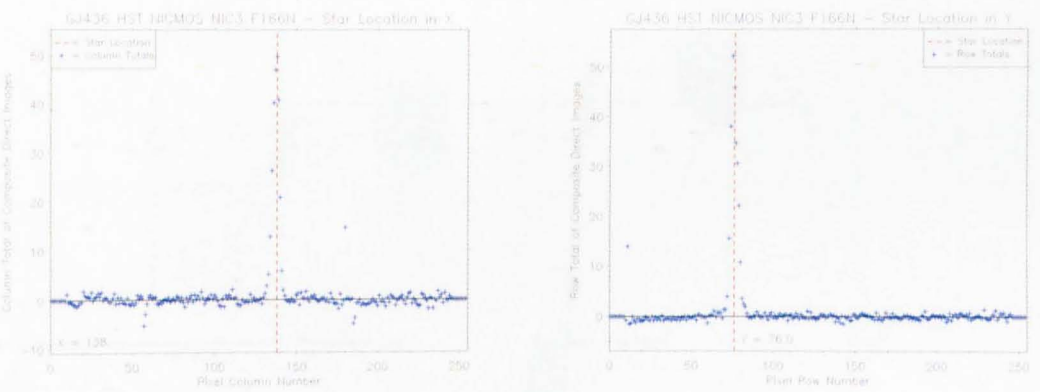
Table 9.2: *Details of the strategy employed for the observations of the transit of GJ436b (upper three sections) and the calculated parameters of the target star and first order spectrum in the resulting images.*

target star in the direct images (see Figure 9.3). After flat-fielding and bad pixel processing, the light in the first order spectrum in each image was gathered via the definition of a light gathering box around said spectrum (see Figure 9.4).

Performing this process on both visits of observation yields the transit light curves presented in Figure 9.5. Note that, unlike previous NICMOS light curves (see Chapters 7 and 8), the transit of GJ436b is much shallower (due to its lesser radius, R_P) and so the signal-to-noise of the transit itself is far less desirable. Thankfully, however, both visits exhibit minimal systematic noise.

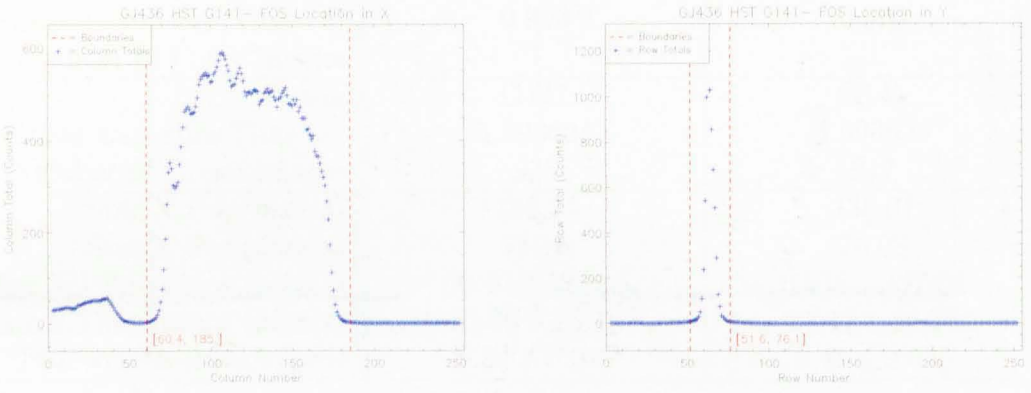


(a) Visit 1

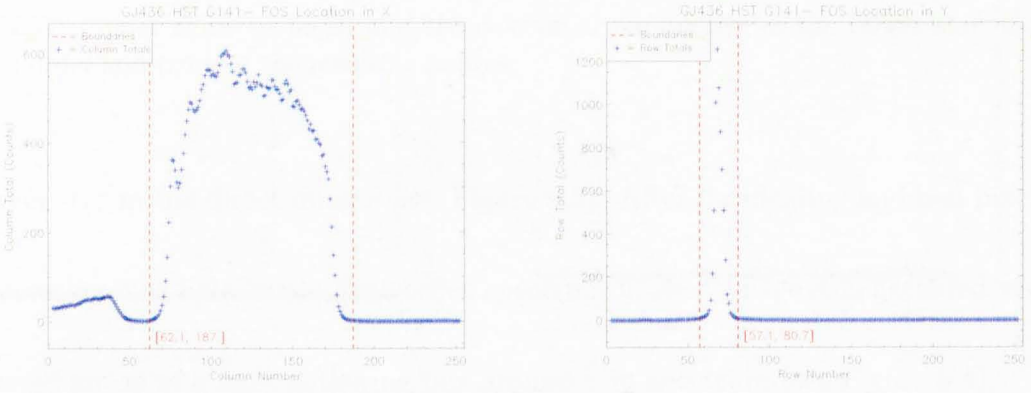


(b) Visit 2

Figure 9.3: *The location (red) of the target star in the filter images of GJ436. Since the data set contains only a single filter image for each visit, this location is calculated from the column- (left) and row-totals (right, blue) of one image rather than a composite image as has previously been the case.*



(a) Visit 1



(b) Visit 2

Figure 9.4: *The dimensions of the light gathering box in the spectral (left) and spatial (right) axes. Each plot features the column- or row-totals (blue) of a composite grism image comprising every exposure in the visit in question.*

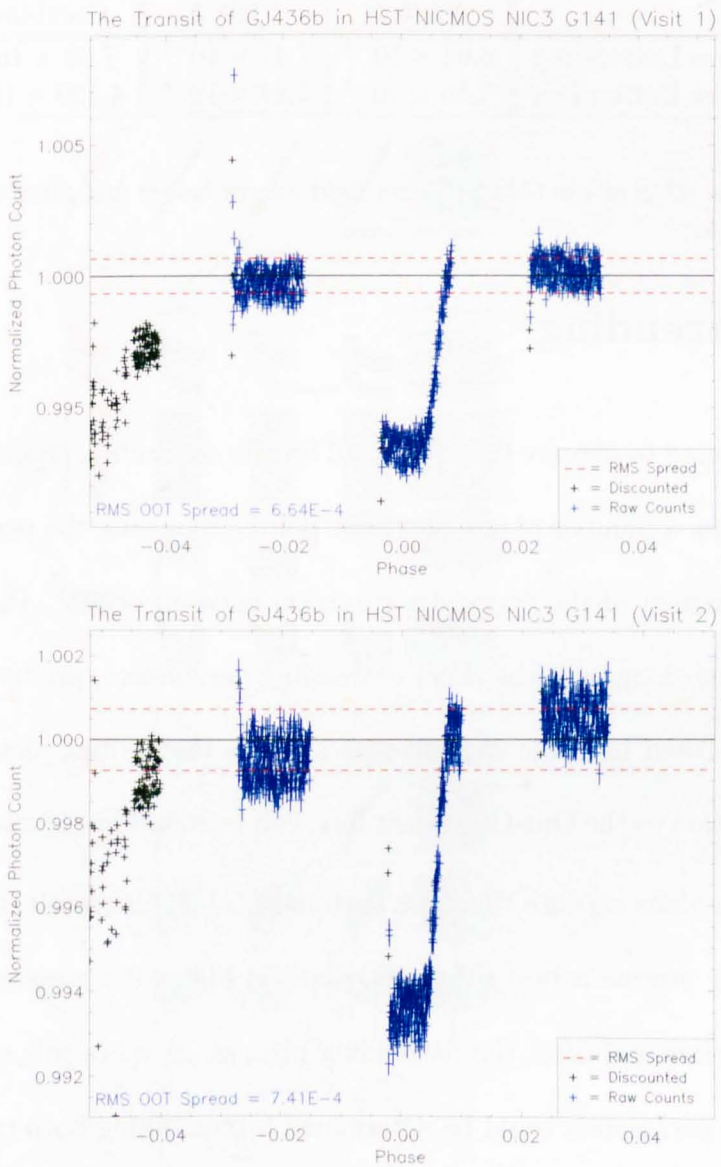


Figure 9.5: *The transit light curves of each visit of the GJ436 data set. Compared to WASP-2b and HD189733b, GJ436b's small radius results in a much shallower transit with an inferior signal-to-noise ratio. It is therefore even more essential that any systematic effects are removed before modeling the transit. Thankfully, the effect of systematic noise in these light curves appears minimal. Note the anomalous points at the beginning of the second orbit; the result of systematics addressed in Section 9.4.*

	Visit 1	Visit 2	Combined
Before Detrending	6.64×10^{-4}	7.41×10^{-4}	7.03×10^{-4}
After Detrending	5.40×10^{-4}	4.82×10^{-4}	4.529×10^{-4}

Table 9.3: *The RMS of the GJ436 transit light curves before and after the removal of systematic noise.*

9.4 Detrending

The strategy used to observe GJ436 is ideal for the detrending process; with each orbit containing a number of out-of-transit points, removing the need to extrapolate the movement of the first order spectrum between orbits[1]. Parameterising this movement, along with the other detrending parameters, produced the RMS improvement listed in Table 9.3. The parameters themselves, along with their modelled relation to the Out-Of-Transit flux, can be found in Figures 9.6 and 9.7.

Since both visits capture the same portion of GJ436b's transit, the impact of the detrending process is best shown in isolation. Figure 9.8 presents each visit's light curve before and after the detrending process. A more robust measure of the planetary parameters could be determined by combining both transits into a single light curve (see Figure 9.9), however, residual systematics in visit 2 means we can only be confident of believable results by modeling the light curve of visit 1 alone.

[1]In orbits containing no out-of-transit points, the detrending model (see Chapter 4) must be extrapolated from other orbits. Even a small number of out-of-transit points in the transit-containing orbit allows a better constraint on the parameter space therein, resulting in a more reliable model.

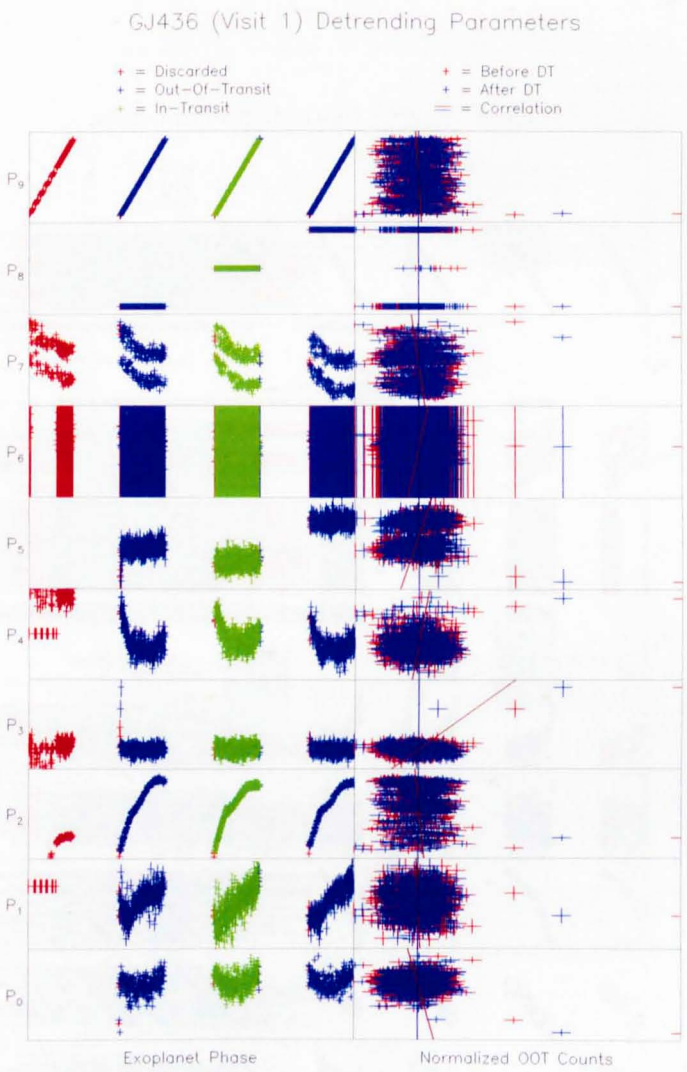


Figure 9.6: *The detrending parameters for visit 1 of the GJ436 data set (left) and their modelled effect on the out-of-transit flux (right). Blue points in the left hand plots are the parameter values for out-of-transit exposures, green points denote in-transit exposures and red points are those exposures removed from processing. The right hand plots show the correlation between out-of-transit flux and the respective parameters both before (red) and after (blue) the detrending process. The bias parameter (P_{10}) showed clear signs of persistence and has been omitted from the detrending process. The anomalous points at the start of the second orbit are reflected in the rotation parameter P_3, but are only partially corrected by the detrending process.*

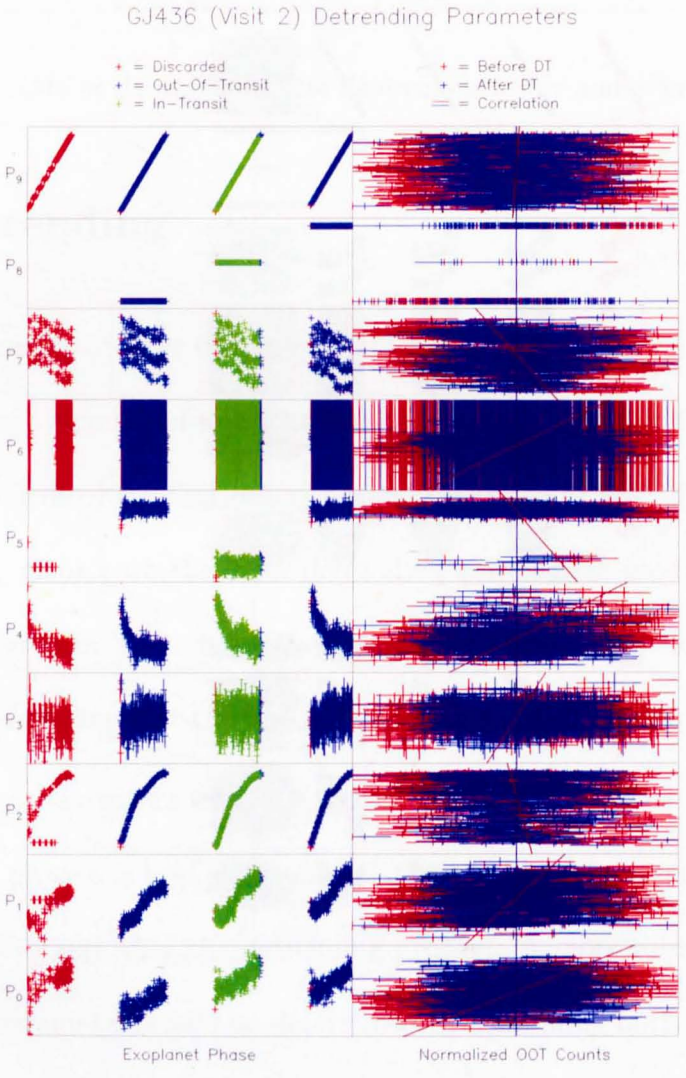


Figure 9.7: *The detrending parameters for visit 2 of the GJ436 data set in the same format as Figure 9.6. Again, the Bias parameter is omited due to persistence. The rotation parameter, P_5, exibits very poor coverage of the in-transit parameter space and may likely result in an imperfect post-detrending light curve.*

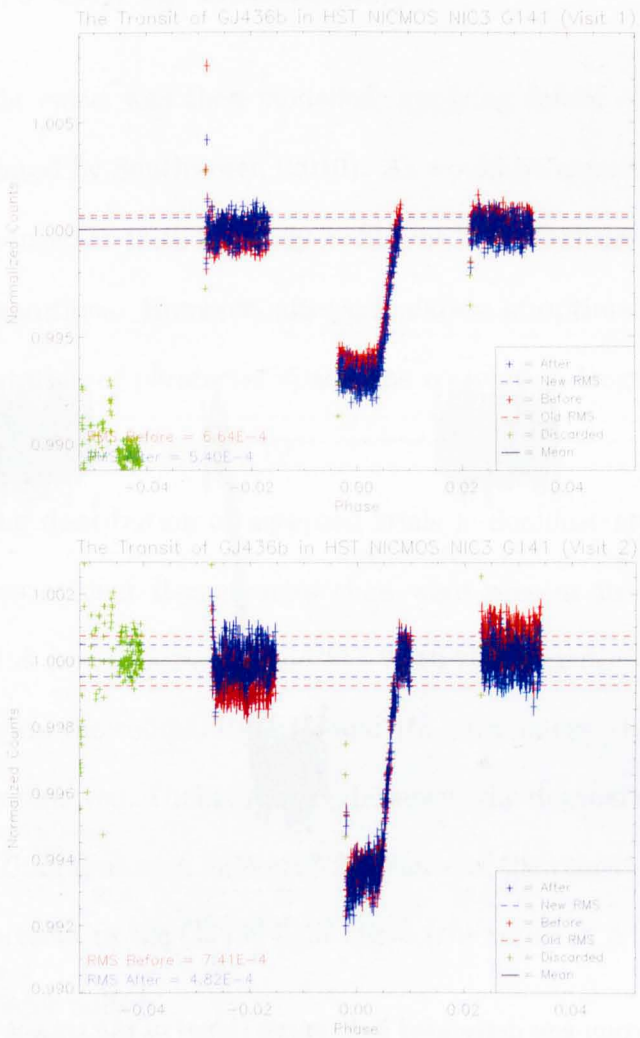


Figure 9.8: *The light curves from visits 1 (top) and 2 (bottom) of the GJ436 data set both before (red) and after (blue) the removal of systematic noise. The dashed lines demonstrate the improvement in the RMS spread of the out-of-transit points as a result of the detrending. Green points are those exposures removed from processing prior to the removal of the systematic effects. Note the clear residual systematic effects in the transit floor of the second visit (orbit 3) - evidence that, as suspected, poor coverage of the in-transit rotation parameter (P_5) has prevented such systematics from being properly removed.*

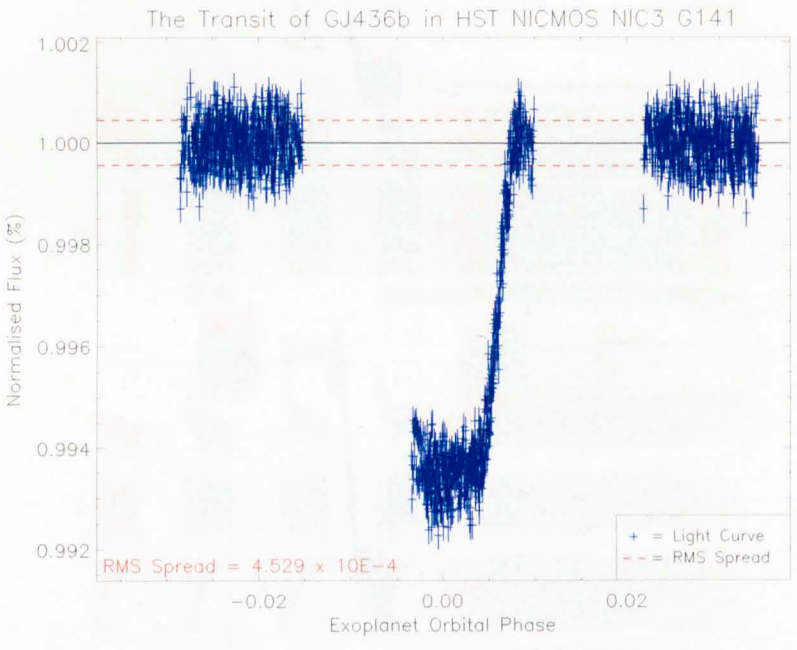


Figure 9.9: *The complete detrended light curve (blue) of the transit of GJ436b across its parent star, comprising both visits of observations. The smooth curve of the visit 1 transit floor is masked by the residual systematics in visit 2. It is therefore best to use simply the visit 1 light curve to model the transit and wavelength-dependent parameters.*

9.5 Light Curve Modeling

The visit 1 light curve was then modelled; applying initial estimates from parameters calculated by Southworth (2010). As would be expected, these starting parameters produced an initial, grossly inefficient MCMC chain with a correlation length of 258 iterations. However, after calculation of optimal step-size weightings and the rotation of parameter space, the correlation length was reduced to 1.22; as demonstrated in Figure 9.10.

The resulting distribution of accepted trials is demonstrated in Figure 9.11 with clearly less evident degeneracies than were present in the WASP-2 and HD189733 chains (see Chapters 7 and 8). With the presence of a transit egress to better constrain the comparative R_S and R_P parameters, the degeneracy seen in HD189733 is removed. Unfortunately, however, the degeneracy between i and a is only lesser than that seen in WASP-2 because of the relatively inferior signal-to-noise ratio present in the GJ436 light curve (the result of a far smaller planet, rather than greater noise).

With the exploration of parameter space properly optimised, iterations continued until the derived ideal solution was found to converge. The histograms found in Figures 9.12 through 9.15 demonstrate the calculation of this ideal solution, together with its related uncertainties, which is listed in Table 9.4.

As is clear from Figure 9.15, despite earlier research suggesting otherwise, there is no evidence of any transit-timing effects in either transit light curve. The values of both ΔT_1 and ΔT_2 are consistent with 0.

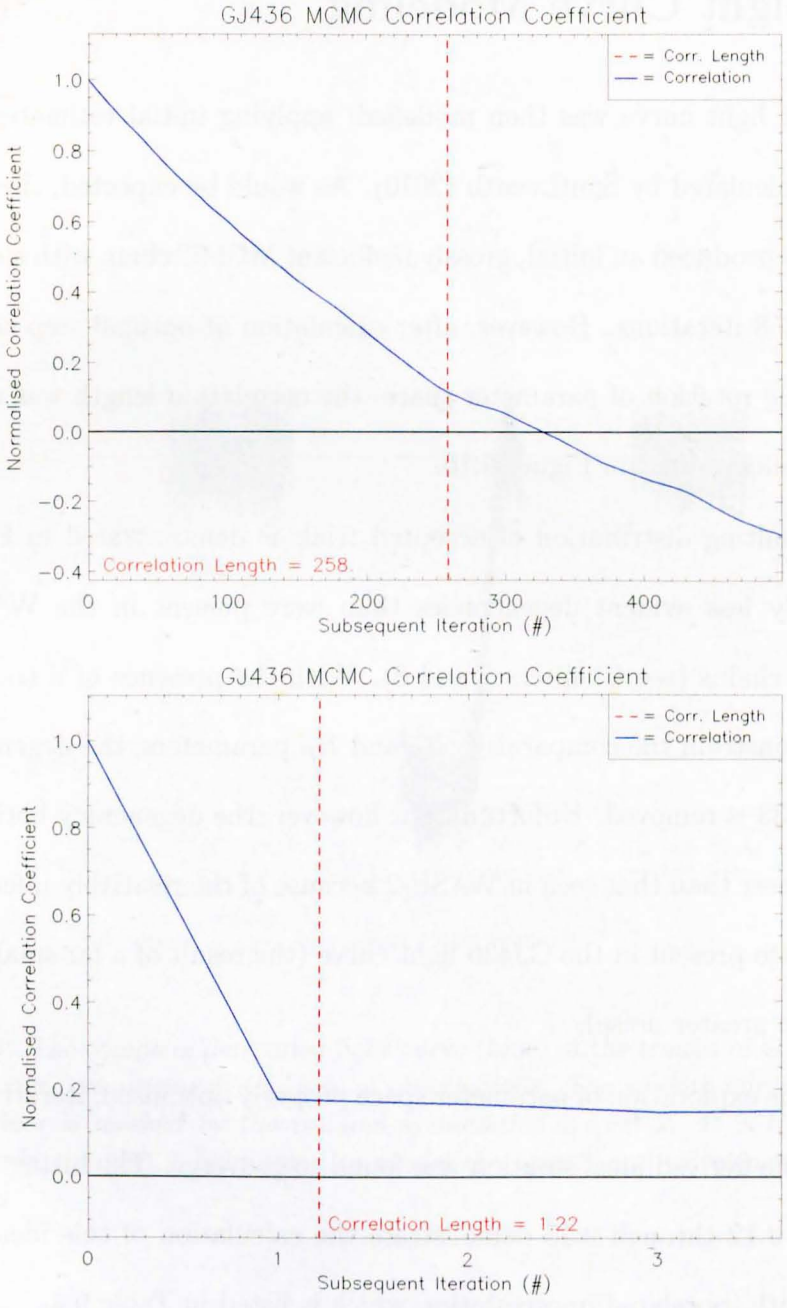


Figure 9.10: *The correlation length of the MCMC simulation at the start of processing (top) and after optimisation (bottom) as detailed in Chapter 5. Note the clear reduction of inheritance in subsequent steps in the chain once the exploration of parameter space has been properly refined.*

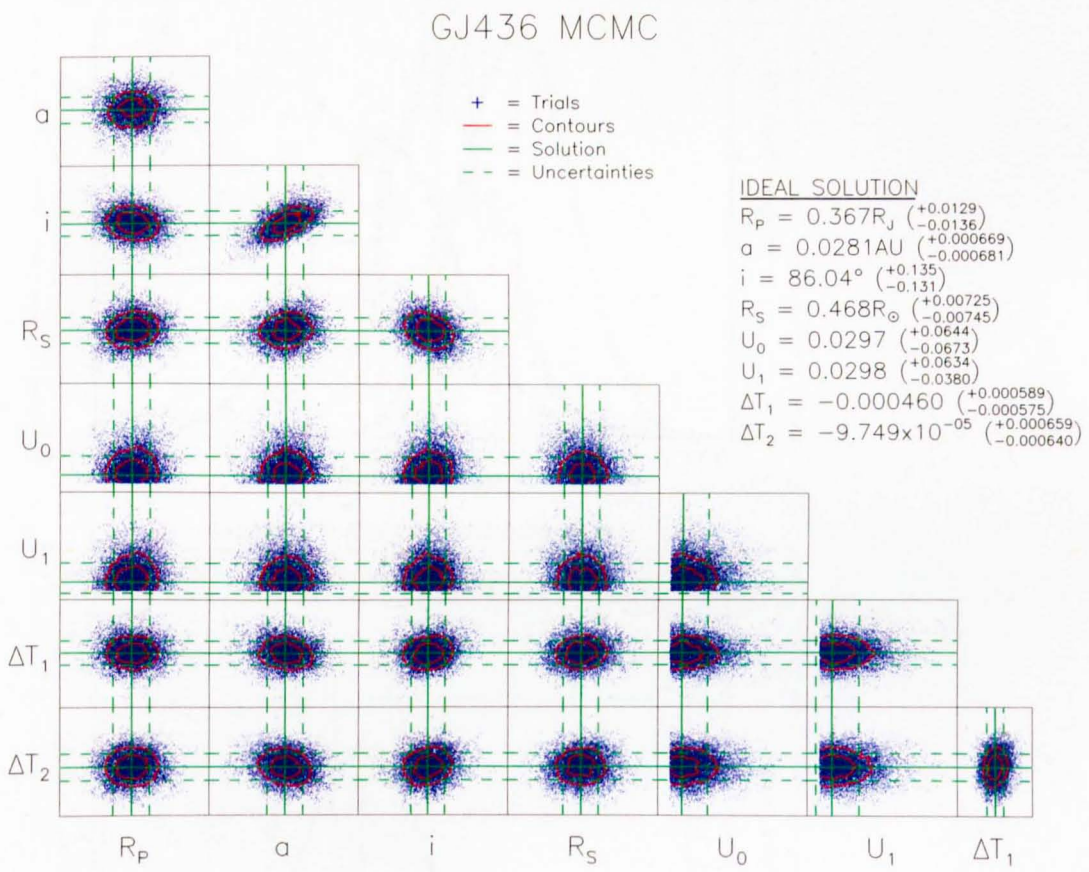


Figure 9.11: *The distribution of accepted trials for a complete MCMC simulation of the GJ436b transit light curve. The overplotted solid green lines show the ideal solution for the parameters calculated from this chain and the associated dashed green lines represent the related uncertainties to that solution. With visit 2 omited from the main detrending process due to its evident systematic errors, the distribution plots of ΔT_2 and the corresponding ideal-solution for that parameter have been calculated from a separate MCMC chain to the other parameters.*

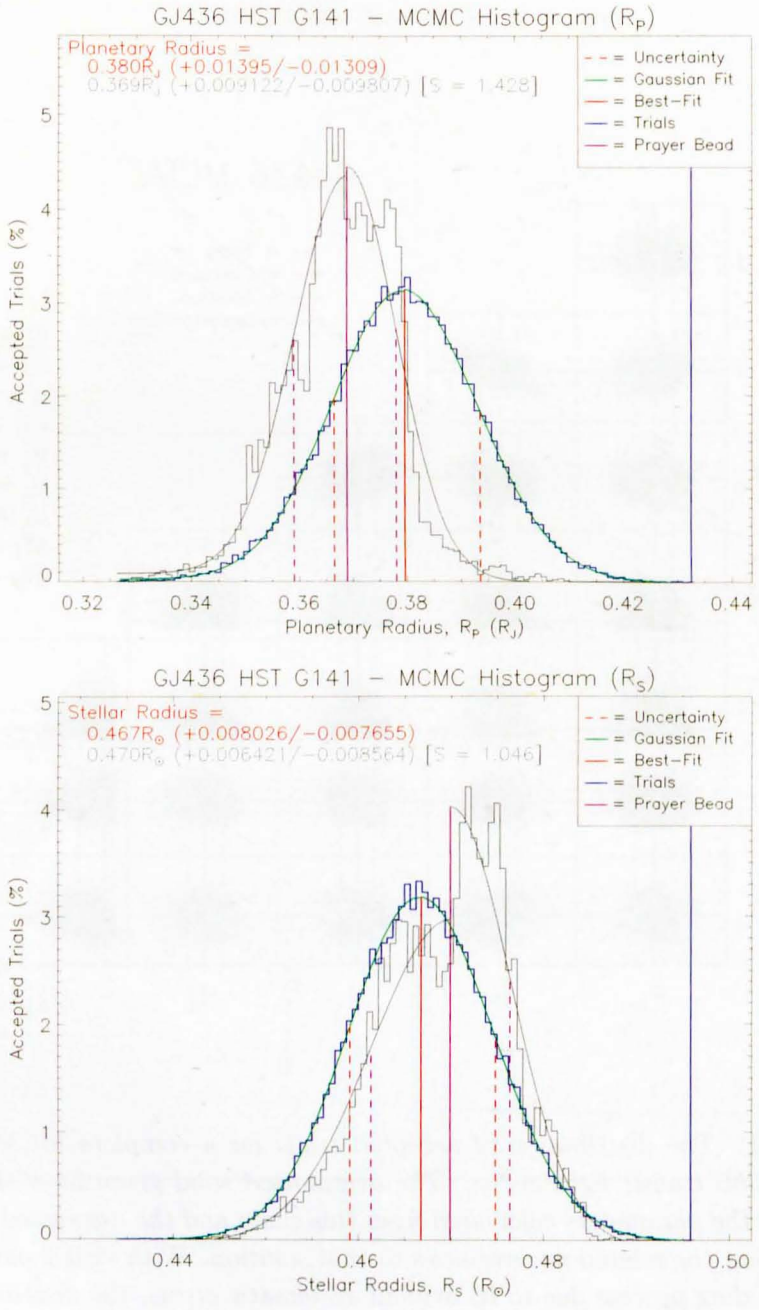


Figure 9.12: *Histograms of accepted MCMC trials in the Planetary and Stellar radius parameters (R_P and R_S) for the GJ436b transit light curve. The ideal solution for these parameters (solid red) and its related uncertainties (dashed red) are calculated from the overplotted assymetric Gaussian curve (green) modelled to the histogram.*

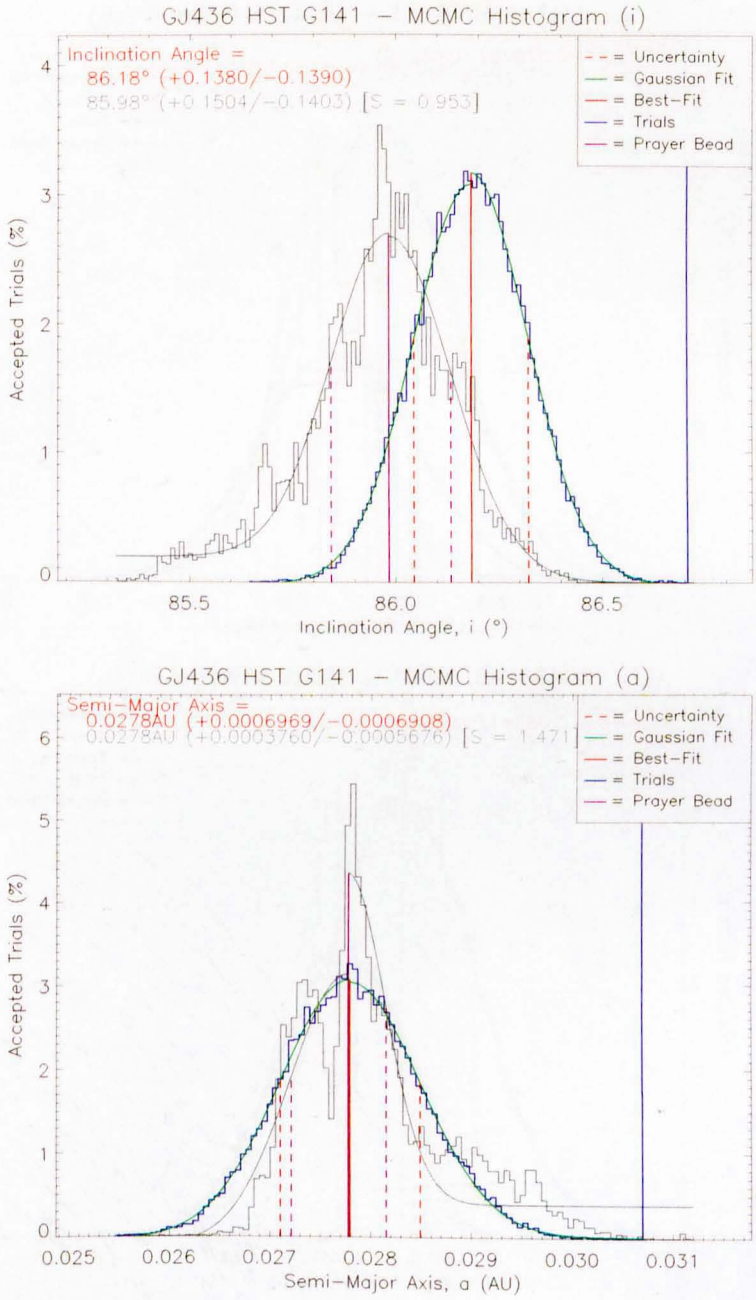


Figure 9.13: *Histograms of accepted MCMC trials in the inclination angle (i) and semi-major axis (a) parameters for the GJ436b transit light curve, presented in the same format employed in Figure 9.12.*

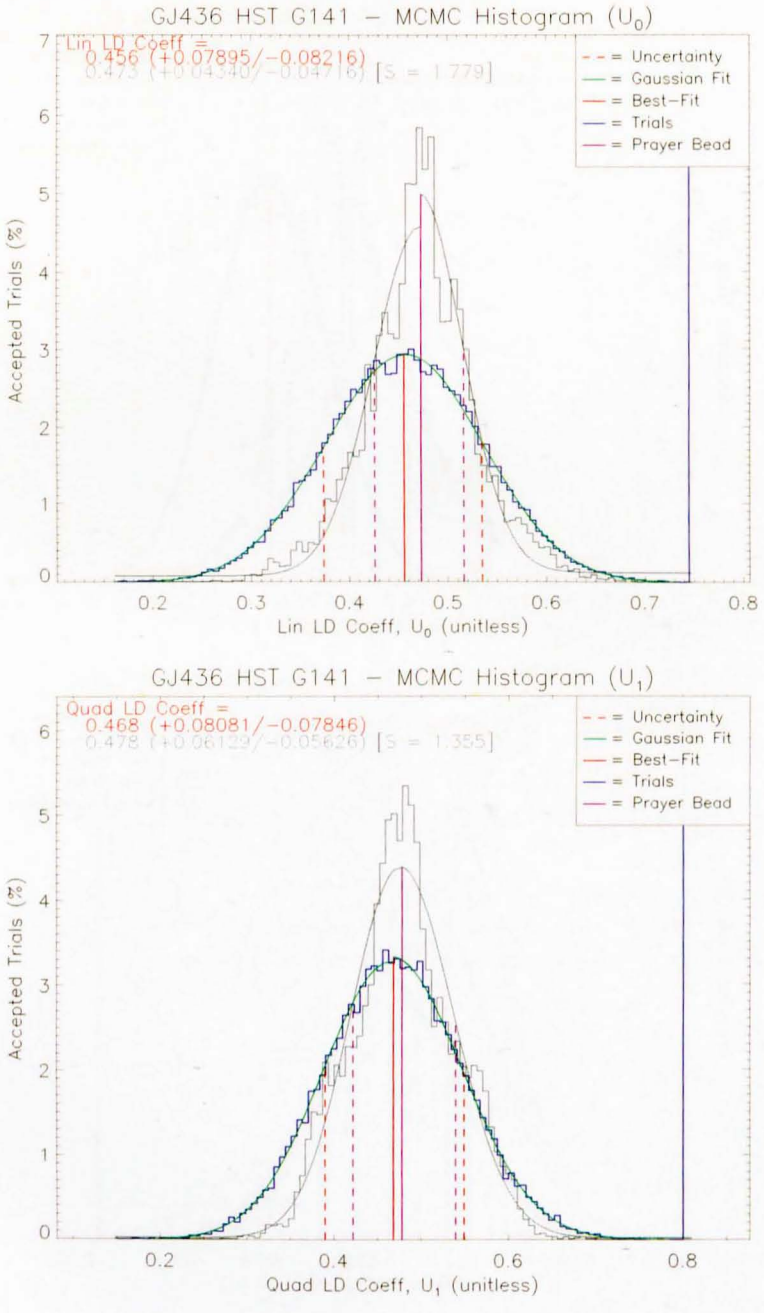


Figure 9.14: *Histograms of accepted MCMC trials in the limb darkening parameters (U_0 and U_1) for the GJ436b transit light curve, presented in the same format employed in Figure 9.12.*

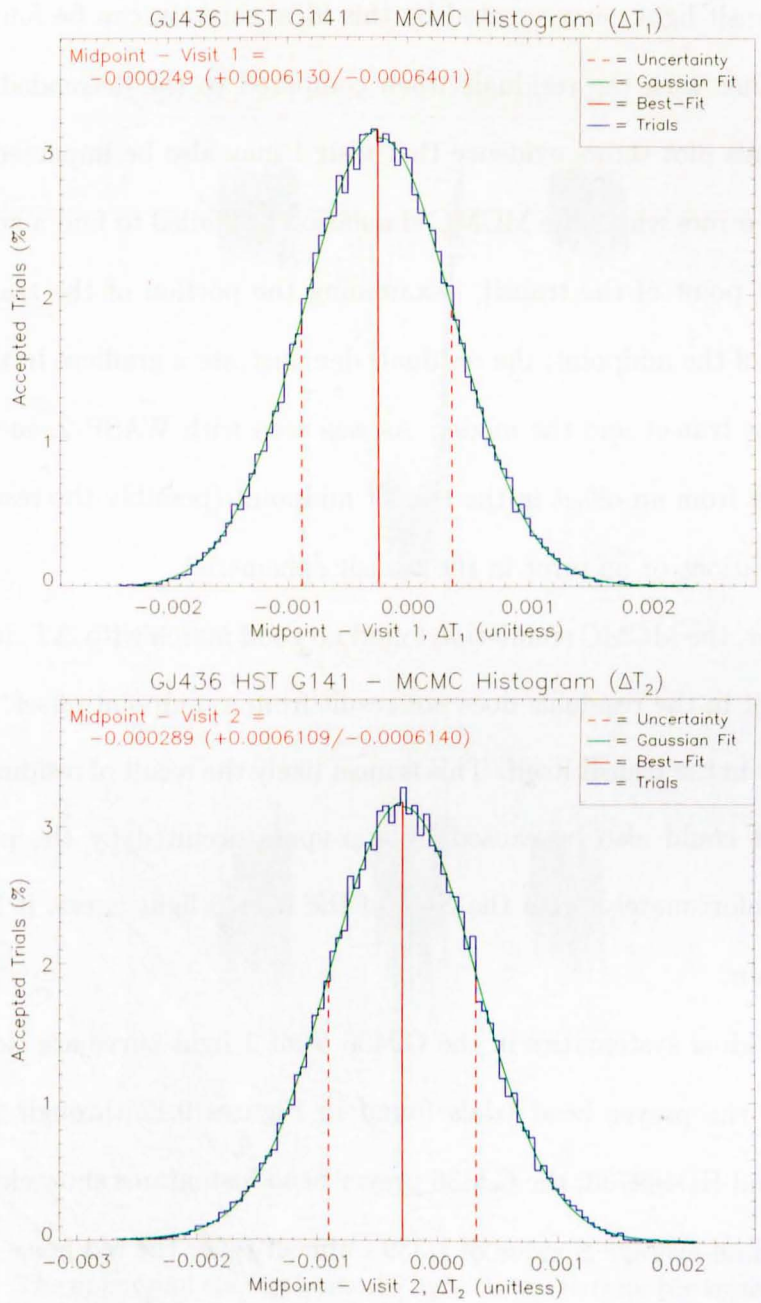


Figure 9.15: *Histograms of accepted MCMC trials in the transit midpoint parameters (ΔT_1 and ΔT_2) for the GJ436b transit light curve, presented in the same format employed in Figure 9.12. Because of the residual systematics in the visit 2 light curve, the value of ΔT_2 has been calculated from an MCMC chain independent of the other parameters.*

The transit light curve created by this ideal solution can be found in Figure 9.16, together with the residuals when compared to the detrended light curve. The residuals plot shows evidence that visit 1 may also be impacted by residual systematic errors where the MCMC simulation has failed to find a good fit to the C3 contact point of the transit. Examining the portion of the transit floor on either side of the midpoint, the residuals demonstrate a gradient in the difference between the transit and the model. As was seen with WASP-2, such a gradient could result from an offset in the transit midpoint (possibly the result of transit timing variations or an error in the transit ephemeris).

However, the MCMC simulation found no good match with ΔT, implying that the gradient in the residuals does not result from a midpoint offset, but from an asymmetry in the transit itself. This is most likely the result of residual systematic effects, but could also be caused by star-spots occulted by the planet during transit. Unfortunately, with the S/N of the GJ436 light curve, it is impossible to be certain.

The residual systematics in the GJ436 Visit 1 light curve are clearly demonstrated by the prayer bead trials found in Figures 9.12 through 9.15. Unlike WASP-2 and HD189733, the GJ436 prayer bead histograms show clear structure, resulting in an average S value of 1.339 - almost twice the red noise remaining in the HD189733 light curve. Such structure is consistent with the Visit 4 feature found in the WASP-2 results (see Chapter 7) - indicating that, despite its superiority to Visit 2, Visit 1 also contains significant residual systematics which throw doubt upon the results. Interestingly, the more dominant systematics are evi-

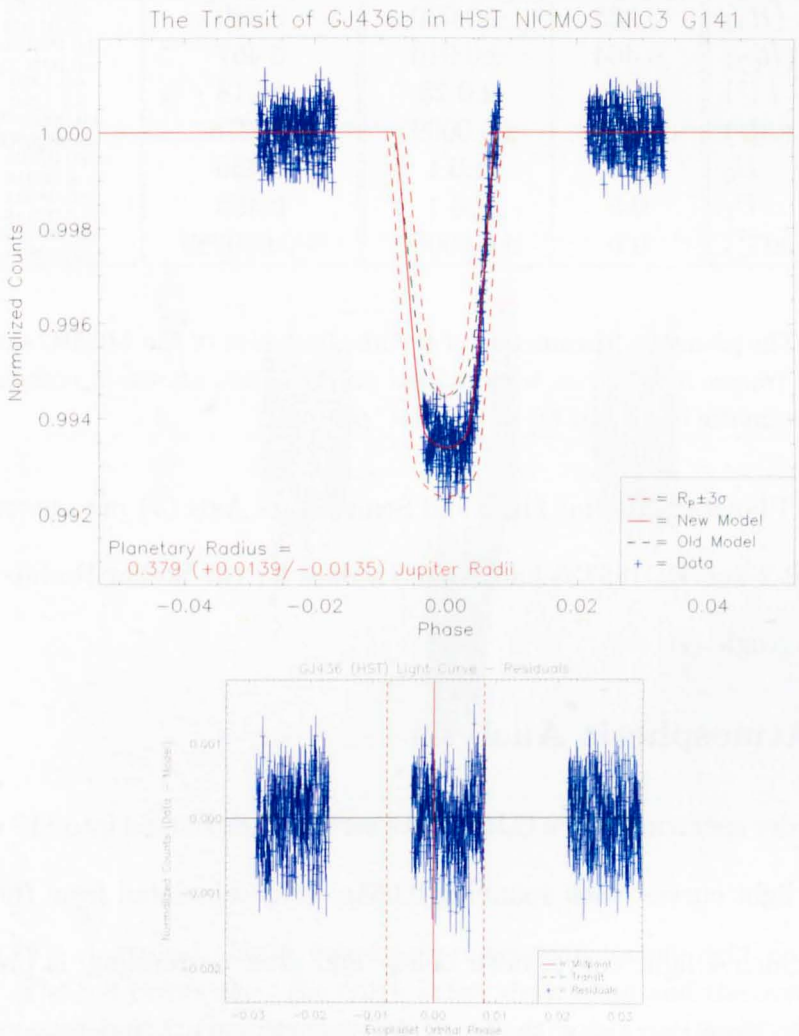


Figure 9.16: *The upper plot shows the model light curve (red) for the transit of GJ436b across its parent star, as derived from the ideal solution of the MCMC simulation of the HST light curve (blue). The lower plot shows the residuals remaining once the model is subtracted from the data (blue) with the transit midpoint and duration denoted by the overploted red lines.*

Parameter	Initial Estimate	Initial Uncertainty	HST G096 Ideal Solution	HST G096 Uncertainty	S Value
R_P (R_J)	0.365	$\pm$0.020	0.380	$^{+0.01395}_{-0.01309}$	1.428
R_S $(R_\odot)$	0.464	$\pm$0.010	0.467	$^{+0.008026}_{-0.007655}$	1.046
i (°)	85.80	$\pm$0.23	86.18	$^{+0.1380}_{-0.1390}$	0.953
a (AU)	0.02887	$\pm$0.00095	0.0278	$^{+0.0006969}_{-0.0006908}$	1.471
U_0	0.5	$\pm$0.1	0.456	$^{+0.07895}_{-0.08216}$	1.779
U_1	0.5	$\pm$0.1	0.468	$^{+0.08081}_{-0.07846}$	1.355
ΔT $(Period^{-1})$	0.0	$\pm$0.00001	$-$0.000289	$^{+0.0006109}_{-0.0006140}$	N/A

Table 9.4: *The planetary parameters of the ideal solution of the MCMC simulation of the GJ436b transit light curve, with related uncertainties, shown in comparison with the initial estimates employed by the MCMC process.*

dent in the Planetary Radius (R_P) and Semi-Major Axis (a) parameters - where both WASP-2 and HD189733 had large S values for the Stellar Radius (R_S) and Inclination Angle (i).

9.5.1 Atmospheric Analysis

The first order spectrum of the GJ436 data set was then binned into 118 wavelenght-dependent light curves, each spanning $0.065\mu m$ (as calculated from the width of the PSF). Such a light curve, both before and after detrending, is presented in Figure 9.17. Note that, as with the overall transit, each λ-dependent curve is a composite of both visits.

Modeling each wavelength-dependent curve in parallel with the main MCMC chain produced a set of model λ-dependent curves presented in Figure 9.18. Converting the results of this model into the relative absorption of the planet produces the absorption spectrum found in Figure 9.19.

From the continuum alone, it is clear that the variations in the transit depth

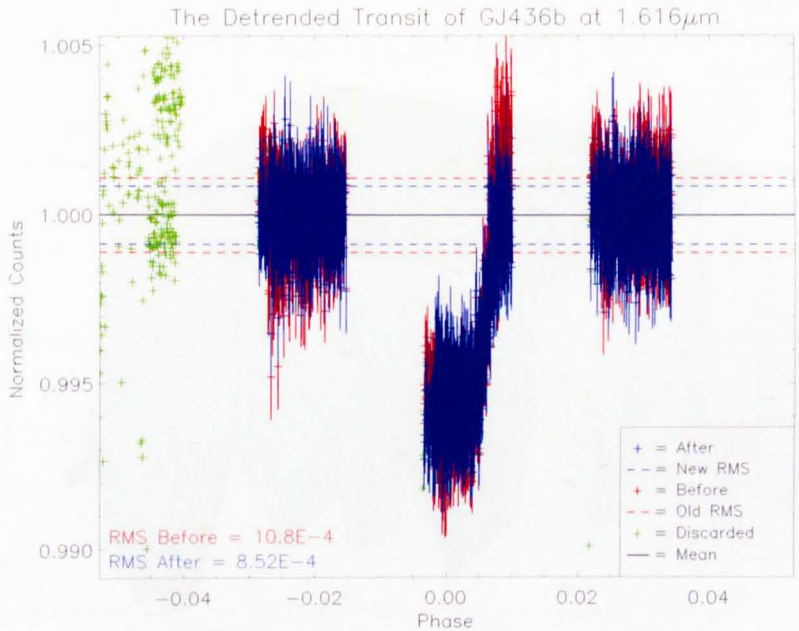


Figure 9.17: *A typical GJ436 wavelength-dependent transit light curve taken at* $1.616\mu m$*. The red points show the curve before detrending and the overplotted blue points the same light curve after detrending. Green points are those values removed from calculation prior to the dentrending process. The poorer photometric precision has greatly degrated the transit when compared to the white light curve (see Figure 9.16), but the transit itself remains distinct.*

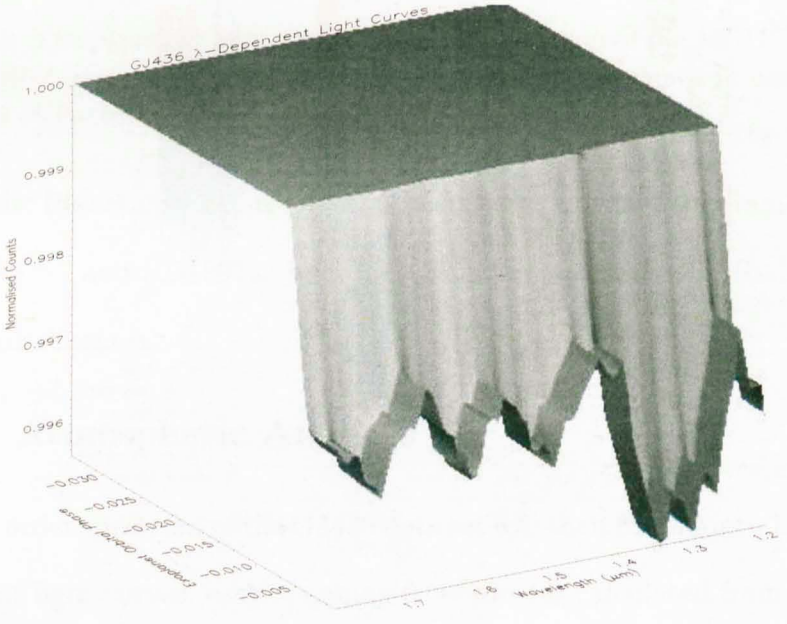


Figure 9.18: *The continuum of models for the wavelength-dependent transit light curves of GJ436b.*

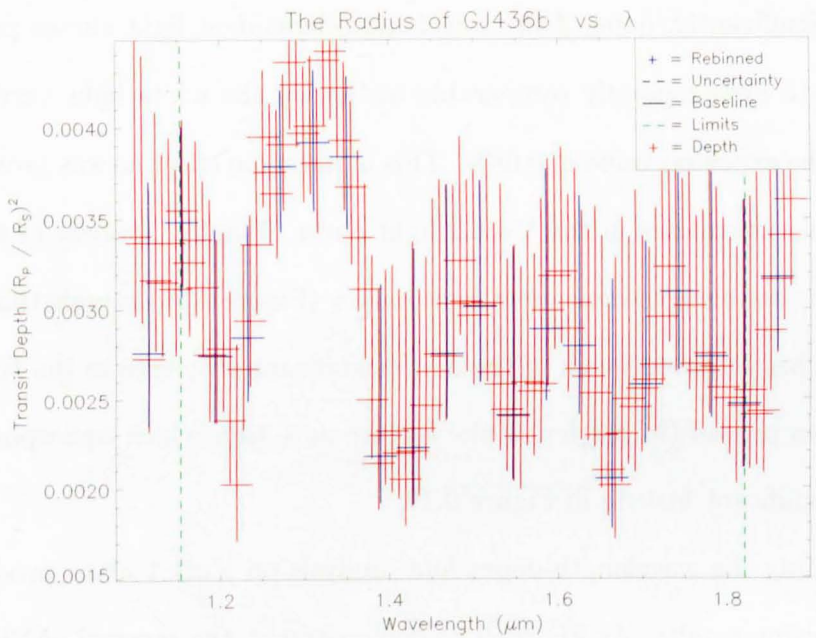


Figure 9.19: *The transmission spectrum of the atmosphere of GJ436b extracted from both visits of observation. The large uncertainties prevent any reliable analysis of the planet's atmosphere, most likely the result of the large systematic effects remaining in the Visit 2 light curve.*

of GJ436b are likely too large to be the result of any atmospheric absorption andare most likely caused by residual systematics - particularly from the Visit 2 light curve. Analysis of the transmission spectrum corroborates this as the large uncertainties on the wavelength-dependent transit depths make all but the feature at $\sim 1.3\mu m$ consistant with each other.

Most significantly, none of the wavelength-dependent light curves produce a transit depth even remotely comparable to that of the white light curve - typically half the expected value of 0.68%. This is the same effect as was produced by the residual systematics in the Visit 2 light curve. Finally, analysis of the RMS values of the wavelength-dependent light curves (Figure 9.20) reveals that - while detrending has removed much of the more significant structure in the RMS, certain features prevail (in particular the feature at $1.3\mu m$ which corresponds with the only significant feature in Figure 9.19.

Performing the wavelength-dependent analysis on Visit 1 alone produces far more promising results. As Figure 9.21 demonstrates, the removal of Visit 2 has produced a much improved set of wavelength-dependent models and hence a far more believable transmission spectrum in Figure 9.22. Furthermore, comparing the RMS plots of the Visit 1 wavelength-dependent curves (Figure 9.23) to those calculated for both visits together (Figure 9.20) shows a significant improvement - both before and after detrending.

The spectrum of GJ436b is difficult to asses as it seems to display two opposing regions. At low wavelengths, the spectrum demonstrates a clear absorption feature at $\sim 1.3\mu m$. This feature exists in the earlier spectrum (see Figure 9.19)

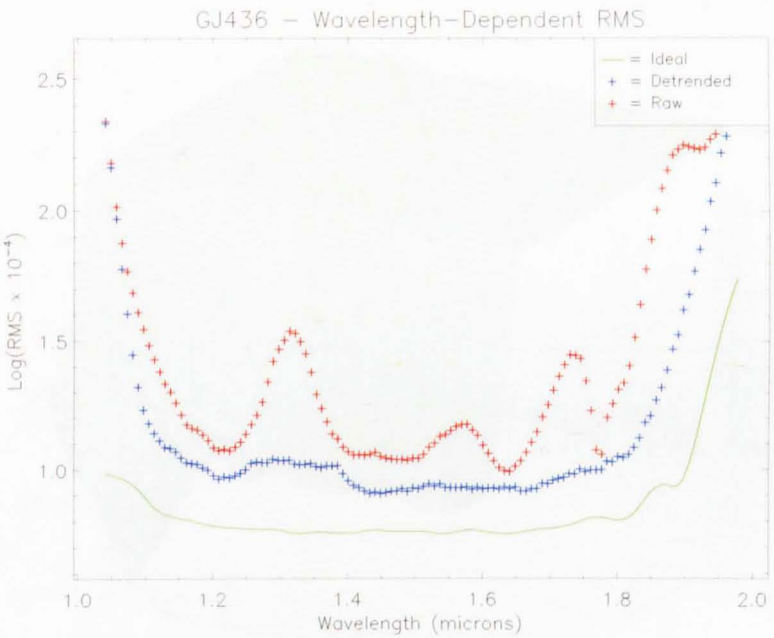


Figure 9.20: *The RMS values of the GJ436 wavelength dependent light curves of both visits calculated before (red) and after (blue) a process of detrending. While the red curve contains many significant features which are removed during detrending, the blue curve still contains a residual sign of the feature at* $1.3\mu m$.

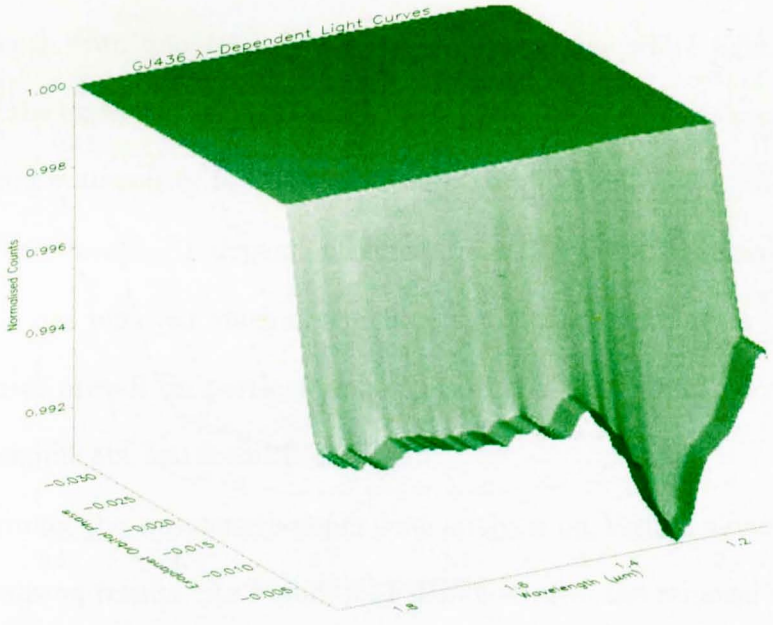


Figure 9.21: *The continuum of models for the wavelength-dependent transit light curves of GJ436b.*

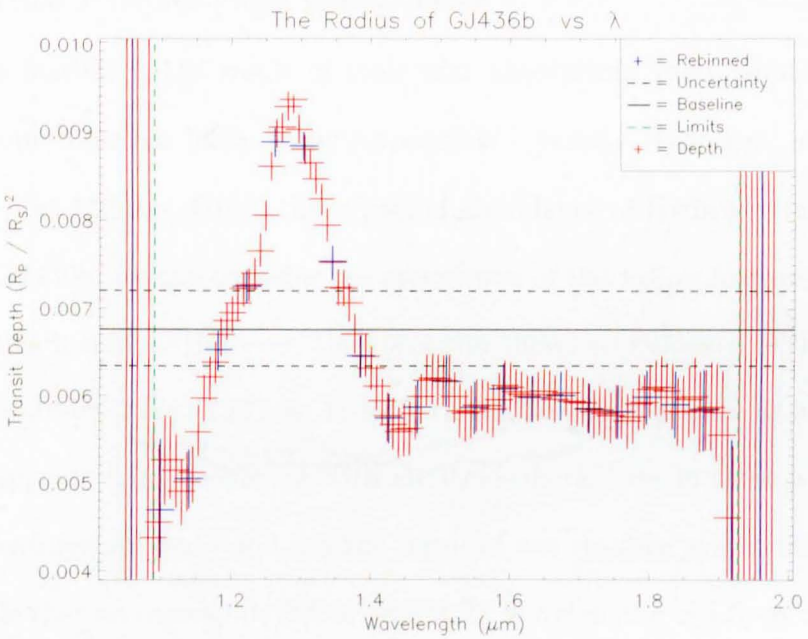


Figure 9.22: *The transmission spectrum of the atmosphere of GJ436b extracted from the Visit 1 data set alone. The feature at* $\sim 1.3\mu m$ *remains with greatly improved uncertainties and a transit depth consistent with the white light curve.*

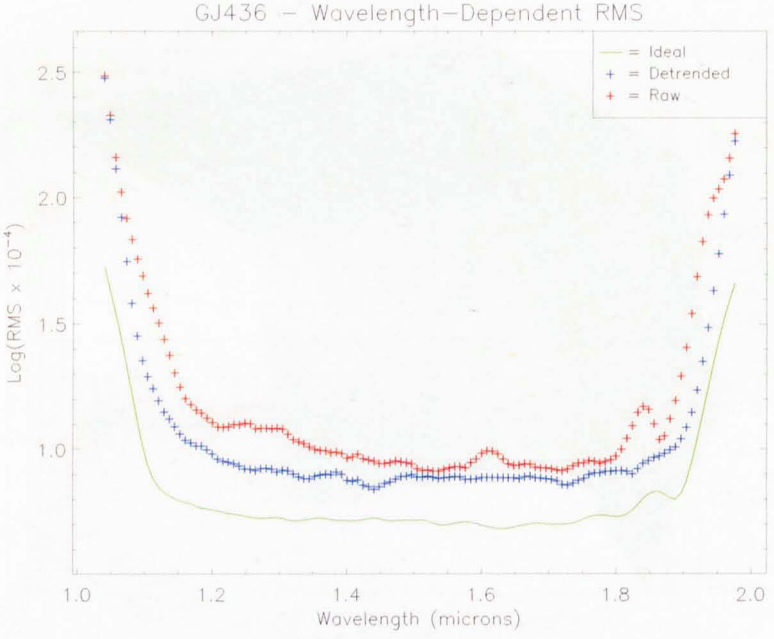


Figure 9.23: *The RMS of the GJ436 Visit* 1 *wavelength-dependent light curves both before (red) and after (blue) a process of detrending. Both curves are a vast improvement on those found in Figure 9.20 and are considerably closer to the calculated optimum RMS spread (green).*

derived from both visits of observation, where it also appeared to correspond with a feature in the RMS Plot (see Figure 9.20). However, the Visit 1 RMS plot does not show such a feature and the $1.3\mu m$ peak is very distinct. The size of the feature also gives us reason to doubt that it could be the result of molecular absorption - 50% deeper than the rest of the spectrum, corresponding to a further 3% reduction in recorded light in-transit.

If the feature is the result of molecular absorption, there would appear to be only one molecule likely to be responsible - namely Hydrogen, absorbing in Paschen-β at $1.28\mu m$. Given the expected abundance of Hydrogen in exoplanets such as GJ436b, we can consider the magnitude of the $1.3\mu m$ feature to be plausible in this instance. However, the spectrum shows no evidence of the expected Pashen-α absorption at $1.87\mu m$. Indeed, the half of the spectrum at higher wavelengths appears flat - consistent with an atmospheric haze in these wavelengths. If both features are real - and not the result of any residual systematics - then it is possible that an atmospheric haze exists for wavelengths $> 1.5\mu m$ with a 'window' around $1.3\mu m$ where evidence of the Paschen-β absorption is not scattered before escaping the atmosphere.

9.6 Conclusions

As with WASP-2 (see Chapter 7), the GJ436 data set provides valuable insight into the reliability of NICMOS data in regard to the systematic effects inherant in such high-precision studies. While the systematics evident in visit 2 could be readily identified from inspection of the light curve (see Figure 9.8), the latent

systematics in visit 1 were not discernable until after the light curve modeling process (see Figure 9.16).

As a result, we must consider the uncertainties calculated by the MCMC process to be an underestimate on the true range of plausible solutions for the light curve model. However, despite this detriment, the parameters and uncertainties of the GJ436 HST NICMOS G141 light curve still represent the best constrained properties of the system calculated to-date.

Despite this, the GJ436 wavelength dependent results have proved very interesting and certainly warrant further research. The possible detection of Paschen-β absorption would be of great significance if it could be confirmed, but with the absence of any feature at $\sim 1.8\mu m$ (for the corresponding Paschen-α absorption) it is impossible to drawn any firm conclusions. Should a means be found to remove the systematic effects plaguing Visit 2, the second transit light curve could either confirm or disprove the veracity of the feature at $1.3\mu m$ - if not determine whether that feature is indeed Paschen-β. For a diffinitive answer, further observations with a different instrument - possibly reaching down to $1.09\mu m$ (Paschen-γ) - would be ideal.

Chapter 10

Conclusions

10.1 Chapter Outline

The following chapter discusses the results of the WASP-2b, HD189733b and GJ436b data sets both in comparison with each other and in regard to the techniques used to derive them. Conclusions are drawn as to the validity and robustness of the programming steps, along with the reliability of the results and what lessons can be learned from the differences between the data sets and their observing strategies. Ultimately, conclusions are drawn as to how the data retrieval and data analysis techniques could be further improved to enhance the precision and believability of future results. The chapter concludes with a review of the future of this field of research and the impact that new techniques and technologies may have on forthcoming projects of this kind.

10.2 Conclusions

10.2.1 Transit Light Curves

As has been clear in all data sets, the greatest obstacle preventing comprehensive analysis is the prevalence of systematic noise in the data and the problems inherent in removing it. While further methods may yet prove successful in improving the techniques detailed in Chapter 4, the best approach would be to take steps to ensure that the data is most simply detrended from the outset. While unapplicable to extant NICMOS data sets (and while the instrument itself is permanently offline), the lessons learned from this and other studies can be applied to future observations with HST's WFC3 (Wide-Field Camera 3) or other such instruments.

In particular, an optimal observing strategy is paramount if the systematic noise is to be best removed. Perhaps paradoxically, as was shown with the WASP-2 data set (see Chapter 7), such a strategy requires not the reduction of systematic noise, but the amplification of it. Dithering, as was employed with WASP-2, has demonstrated how a more comprehensive model of the positional systematics may be achieved. As was noted in Section 7.7, a similar approach could be applied to many of the parameters used to constrain the position of the first order spectrum (provided that the dithering only occurs in out-of-transit exposures).

Such an approach can more readily assure a good coverage of parameter space - the lack of which proved costly for visit 4 of the WASP-2 data set and both visits of observation to GJ436 (see Chapter 9). As was also demonstrated by visit

4 of the WASP-2 data, long term trends (as were particularly evident in WASP-2 visit 3 and the secondary eclipse of HD189733b, see Chapter 8) can only be properly constrained with multiple out-of-transit orbits. Without this approach, long term trends become degenerate with trends affecting a single orbit (such as the trends in the primary transit of HD189733b).

Furthermore, as was noted in analysis of GJ436, detrending is more reliable when in-transit orbits contain some portion of out-of-transit points to better constrain their parameter space during the detrending process. While this may prevent observation of a complete transit light curve, the MCMC modeling procedure detailed in Chapter 5 does not require a complete light curve to determine accurate parameters. Indeed, as was shown in the HD189733 data set, lack of coverage of the transit ingress and egress can have a greater detriment to the precision of the derived parameters than only a partial coverage of the transit floor.

In summary, an ideal visit of observation to observe the transit of an exoplanet by high-precision grism spectroscopy (such as with NICMOS or WFC3) should adopt the following points in its observing strategy:

1. At least four HST orbits: The first to be discarded while the telescope is allowed to settle into its new pointing; two of the remaining three to capture only out-of-transit light, thereby providing long-term coverage of parameter space for future systematic noise reduction; and one in-transit orbit, preferably situated between the two out-of-transit orbits to allow

interpolation of any trend between the two rather than a less precise process of extrapolation.

2. A process of dithering adopted for at least one out of transit orbit and in multiple parameters to give as good a coverage of the systematic parameter space as possible.

3. At least 10 out-of-transit exposures conducted in the in-transit orbit to better constrain its position in parameter space. Depending on the duration of transit, it may not be possible to capture any of the transit floor with such an approach, in this case, the most preferable solution would be for a second in-transit orbit capturing some portion of the transit floor with the base of the ingress/egress while the other in-transit orbit captures the top of the egress/ingress and a portion of out-of-transit points.

In terms of the method of detrending, Preferential Decorrelation appears to have proven itself superior to the traditional Linear Regression method of removing systematic noise. As Table 10.1 shows, the Preferential method of detrending appears to produce a model which extrapolates more reliably into the in-transit parameter space with superior S values in all three data sets (though for HD189733, the data set least effected by systematic noise, there is almost no difference).

System	Visit Number	Raw RMS	Preferential Decorrelation RMS	Mean S Value	Linear Regression RMS	Mean S Value
WASP-2	3	13.5	5.29	1.214	5.55	1.230
	4	12.4	4.71	0.770	4.74	0.506
HD189733	1	5.63	4.11	1.119	4.34	1.207
	2	5.87	4.74	1.204	4.74	1.212
GJ436	1	5.43	4.44	1.339	6.07	1.406
	2	7.30	4.73	1.553	5.35	1.583

Table 10.1: *RMS Values for each visit of each data set as processed by the Preferential Decorreltion and Linear Regression detrending techniques, in reference to the S value measuring the residual systematic noise in each case. (All values given are 10^{-4}.)*

10.2.2 Atmospheric Spectra

As has been demonstrated in every data set, the prevalence of systematic noise in such high-precision observations has made any firm conclusions difficult to draw. It should be noted, however, that NICMOS (together with the appropriate data reduction and light curve modeling tools) has proved itself capable of detecting very small changes in transit depth.

With the development of comparative methods to gauge the impact of systematic noise on the results[1], it has become easier to determine some of the impact of residual systematic noise on results. However, with the true degree of residual systematics impossible to measure, it is impossible to say what proportion of systematics can be identified by such techniques.

The best approach would seem to be for follow-up observations to be con-

[1]More specfically: the inspection of the post-detrending RMS spreads of the wavelength-dependent light curves, in relation to the derived values of $R_P(\lambda)$; and the application of overlapping bins allowing identification of variations in $R_P(\lambda)$ with lesser width than the PSF.

ducted by a variety of different instruments across the wavelength ranges of the NICMOS grisms and the data sets in question. Once a variety of observations agree on the true transmission spectrum of an exoplanets, those data sets or analyses which produced false results can then be assessed to determine the source of their inaccuracy and help prevent it from influencing future studies.

Setting aside concerns of the results' believability, analysis of the extracted spectra demonstrates no prevalence for either atmospheric hazes or molecular absorption features. While WASP-2b may demonstrate absorption features consistent with water (see Figure 7.26), the apparent slant of the spetrum (if real) could only be the result of such a haze. Similarly GJ436b, with its large unexplained absorption at $1.3\mu m$, also presents a consitently flat continuum upwards of $1.4\mu m$ consistent with a haze. Only HD189733b demonstrates distinct absorption features (and, potentially, emission features); but this is contrary to previous studies which have concluded that the transmission spectrum of HD189733b demonstrates only an atmospheric haze.

10.3 Further Research

As with any data-reduction pipeline, there are always ways in which the program can be improved or refined. While the endeavour of this research has been to exhaust the potential scientific output of each data set, there are still advances that could be made in both the efficiency and scope of the pipeline as a whole.

Naturally, all research presented herein is robust, precise and as detailed as was possible over the timescale of the research. However, that is not to say that

further developments in the pipeline, or the field as a whole, may not yet find the means to eke more results from HST transit grism spectroscopy.

10.3.1 Limb Darkening

As was evident in the results of HD189733 (see Chapter 8) the means by which stellar limb darkening is modelled can have a significant impact on the measured $(R_P/R_S)^2$ of the transit light curve. For planets with well-constrained parameters, transit light curves may also prove a means by which to study the limb darkening of stars and arrive at a more comprehensive model for use in future studies.

10.3.2 MCMC Detrending

The most significant potential improvement to the existing pipeline would be the integration of the established detrending process (see Chapter 4) with the MCMC light curve modeling (see Chapter 5). Combining these two processes would allow both the detrending coefficients and the model light curve parameters to be calculated simultaneously; potentially improving both the removal of HST's systematic effects and the accuracy of the resulting model parameters.

Any improvement would result from the integrated pipeline's ability to model any in-transit systematic effects directly, without having to extrapolate from the Out-Of-Transit (OOT) points. In this approach, the light curve systematics can be modelled from the residuals $(L - \Gamma)$ of the current model light curve and so, as the model converges on an ideal solution, the detrending process will also converge on an ideal detrended light curve against which that model can be compared.

10.3.3 Wavelength Deconvolution

Another potential improvement to the data-processing pipeline would be the application of a deconvolution algorithm to reduce or remove the effects of defocussing on the resulting atmospheric spectra.

Currently, light falling on the detector at any given wavelength will be distributed according to the point spread function (or PSF) of the telescope's optics. In the case of defocussed exposures, the PSF is a large ($\sim$ 10-15 pixels in diameter), blurred annulus. The resulting first order spectrum on the detector can be considered a continuum of overlapping annuli, each representing light at a different wavelength.

The exact shape of the PSF can be modelled from the direct images taken as part of each data set. It may be possible to deconvolve such a model the first order spectrum and thus recover the wavelength information lost during defocussing. Correct application of such an approach could, theoretically, improve the wavelength resolution of the result to a limit governed only by the width of the detector's pixels.

10.3.4 Optimum Resolution

Comparing any overall light curve with a wavelength-dependent light curve from the same data set, it is evident that the overall light curves represent a relatively small improvement for their far superior (typically $> 10\times$) light-gathering region. This demonstrates, as would be expected, that improving the light-gathering region for any given light curve yields only diminishing returns in the resulting

precision.

Typically, such an increase in light-gathering area would result from an increase in the defocussing applied by the obverving strategy. However, any such increase in defocussing would also result in a reduction in the wavelength-resolution of the resulting spectrum as light from a wider range of wavelengths would fall in the same pixels.

In theory, there should be some optimum level of defocussing whereby the highest precision spectra can be extracted from the observations. An analysis across all NICMOS observations, potentially coupled with a model of the detector itself, should be capable of determining the best degree of defocussing to produce wavelength-dependent light curves with the smallest possible RMS Spread compared to the greatest possible wavelength-resolution.

Unfortunately, since NICMOS is no longer in operation, this optimum observing strategy could never be put into practice. However, it may prove informative when devising future observing strategies for other instruments performing grism spectroscopy on exoplanetary atmospheres.

10.3.5 Future Instruments

Unfortunately, NICMOS, the instrument responsible for the data reported in this thesis, went offline in 2008 after a series of technical failures. With Hubble's operational lifespan nearing its end, it is unlikely that any attempt to revive NICMOS will ever be executed. However, identical (if not superior) grism spectroscopy can already be performed using Hubble's newly-installed Wide-Field

Camera 3 (WFC3), which has already been applied to observe the transit of exoplanet WASP-3b. It is likely that WFC3 will succeed NICMOS at the forefront of exoplanetary research and should be expected to produce the most detailed exoplanetary atmospheric spectra to-date over the remainder of HST's time in service.

Meanwhile, ground-based observations continue to apply new technologies and methods to close the gap on their space-based rivals. As systems such as adaptive optics become more advanced and widespread, it is probable that ground based observations will begin to challenge those of Hubble and Spitzer - though, of course, space-based instruments will remain superior, especially at infrared wavelengths where ground-based observations are limited by the Earth's atmosphere.

Looking still further ahead, the James Webb Space Telescope (JWST, Hubble's successor) should produce results eclipsing all those which have come before. JWST's NIRSpec instrument will be capable of grism spectroscopy, much like NICMOS and WFC3, but with far greater precision and resolution. With the right observing technique and proper data analysis, JWST is expected to have the capability to observe, not just the transmission spectra of Hot Jupiters, but the reflected spectra of terrestrial exoplanets during secondary eclipse.

With the incredible success of the Kepler space telescope - which is currently working to confirm the detection of over one-thousand transiting exoplanets - it is likely that there will be no shortage of targets for JWST after it begins operation (the satelite is currently scheduled to launch in June 2014). With the right target exoplanet and the right application of JWST, it may even be possible to detect

compounds such as Oxygen (O_2) or Ozone (O_3) on another world - producing (if it is there to be found) the most conclusive proof to-date of the existance of extraterrestrial life.

Following on from Kepler, two missions currently competing for funding in the ESA Cosmic Visions program could advance the field still further. PLATO (PLAnetary Transits and Oscillations of stars) will conduct a multi-camera, wide-field transit survey, modelled on the successful SuperWASP ground-based telescopes. In conjunction, ECHO (Exoplanet CHaracterisation Observatory), will be the first instrument designed bespokely to observe the atmospheres of exoplanets and will surpass even JWST in the molecules it may one day discover.

10.4 Closing Remarks

The results and conclusions presented in this thesis are simply the first steps in the new and exciting field of exoplanetary atmospheric physics. As research continues, it can be hoped that observations succeeding those detailed herein will provide some of the most exciting and, in all likelihood, unexpected results since the discovery of the first exoplanet itself - addressing such unanswered questions of the processes of planet formation and the search for life elsewhere in the universe.

References and Bibliography

Y. Alibert, C. Mordasini, and W. Benz. Extrasolar planet population synthesis III. Formation of planets around stars with different masses. *A&A*, Awaiting Publication, 2010. 2

E. K. Baines, H. A. McAlister, T. A. ten Brummelaar, J. Sturmann, L. Sturmann, N. H. Turner, and S. T. Ridgway. Eleven Exoplanet Host Star Angular Diameters from the Chara Array. *ApJ*, 701:154–162, August 2009. doi: 10.1088/0004-637X/701/1/154. 260

J. R. Barnes, T. S. Barman, H. R. A. Jones, R. J. Barber, B. M. S. Hansen, L. Prato, E. L. Rice, C. J. Leigh, A. Collier Cameron, and D. J. Pinfield. A search for molecules in the atmosphere of HD 189733b. *MNRAS*, 401:445–454, January 2010. doi: 10.1111/j.1365-2966.2009.15654.x. 263

J. L. Bean, G. F. Benedict, D. Charbonneau, D. Homeier, D. C. Taylor, B. McArthur, A. Seifahrt, S. Dreizler, and A. Reiners. A Hubble Space Telescope transit light curve for GJ 436b. *A&A*, 486:1039–1046, August 2008. doi: 10.1051/0004-6361:200810013. 302

J. P. Beaulieu, D. M. Kipping, V. Batista, G. Tinetti, I. Ribas, S. Carey, J. A. Noriega-Crespo, C. A. Griffith, G. Campanella, S. Dong, J. Tennyson, R. J. Barber, P. Deroo, S. J. Fossey, D. Liang, M. R. Swain, Y. Yung, and N. Allard. Water in the atmosphere of HD209458b from 3.6-8μm IRAC photometric observations in primary transit. *MNRAS*, Awaiting Publication, 2010. 45

D. P. Bennett, I. A. Bond, A. Udalski, T. Sumi, F. Abe1, A. Fukui, K. Furusawa, J. B. Hearnshaw, S. Holderness, Y. Itow, K. Kamiya, A. V. Korpela, P. M. Kilmartin, W. Lin, C. H. Ling, K. Masuda, Y. Matsubara, N. Miyake, Y. Muraki, M. Nagaya, T. Okumura, K. Ohnishi, Y. C. Perrott, N. J. Rattenbury, T. Sako,

To. Saito, S. Sato, L. Skuljan, D. J. Sullivan, W. L. Sweatman, P. J. Tristram, P. C. M. Yock, M. Kubiak, M. K. Szymański, G. Pietrzyński, I. Soszyński, O. Szewczyk, L. Wyrzykowski, K. Ulaczyk, V. Batista, J. P. Beaulieu, S. Brillant, A. Cassan, P. Fouqué, P. Kervella, D. Kubas, and J. B. Marquette. A low-mass planet with a possible sub-stellar-mass host in microlensing event MOA-2007BLG-192. *ApJ*, 684:663–683, 2008. 36

I. A. Bond, A. Udalski, M. Jaroszyński, N. J. Rattenbury, B. Paczyński, I. Soszyński, L. Wyrzykowski, M. K. Szymański, M. Kubiak, O. Szewczyk, K. Żebruń, G. Pietrzyński, F. Abe, D. P. Bennett, S. Eguchi, Y. Furuta, J. B. Hearnshaw, K. Kamiya, P. M. Kilmartin, Y. Kurata, K. Masuda, Y. Matsubara, Y. Muraki, S. Noda, K. Okajima, T. Sako, T. Sekiguchi, D. J. Sullivan, T. Sumi, P. J. Tristram, T. Yanagisawa, P. C. M. Yock, The MOA, and OGLE Collaborations. OGLE-2003-BLG-235/MOA 2003-BLG-53: A planetary microlensing event. *ApJ*, 606:L155–L158, 2004. 34, 36

W. J. Borucki, D. Koch, G. Basri, N. Batalha, T. Brown, D. Caldwell, J. Caldwell, J. Christensen-Dalsgaard, W. D. Cochran, E. DeVore, E. W. Dunham, A. K. Dupree, T. N. Gautier III, J. C. Geary, R. Gilliland, A. Gould, S. B. Howell, J. M. Jenkins, Y. Kondo, D. W. Latham, G. W. Marcy, S. Meibom, H. Kjeldsen, J. J. Lissauer, D. G. Monet, D. Morrison, D. Sasselov, J. Tarter, A. Boss, D. Brownlee, T. Owen, D. Buzasi, D. Charbonneau, L. Doyle, J. Fortney, E. B. Ford, M. J. Holman, S. Seager, J. H. Steffen, W. F. Welsh, J. Rowe, H. Anderson, L. Buchhave, D. Ciardi, L. Walkowicz, W. Sherry, E. Horch, H. Isaacson, M. E. Everett, D. Fischer, G. Torres, J. A. Johnson, M. End, P. MacQueen, S. T. Bryson, J. Dotson, M. Haas, J. Kolodziejczak, J. van Cleve, H. Chandrasekaran, J. D. Twicken, E. V. Quintana, B.D. Clarke, C. Allen, J. Li, H. Wu, P. Tenenbaum, E. Verner, F. Bruhweiler, J. Barnes, and A. Prsa. Kepler planet detection mission: introduction and first results. *Science*, 327: 977–980, 2010. 46

W. J. Borucki, D. G. Koch, G. Basri, N. Batalha, T. M. Brown, S. T. Bryson, D. Caldwell, J. Christensen-Dalsgaard, W. D. Cochran, E. DeVore, E. W. Dunham, T. N. Gautier, III, J. C. Geary, R. Gilliland, A. Gould, S. B. Howell, J. M. Jenkins, D. W. Latham, J. J. Lissauer, G. W. Marcy, J. Rowe, D. Sasselov, A. Boss, D. Charbonneau, D. Ciardi, L. Doyle, A. K. Dupree, E. B. Ford,

J. Fortney, M. J. Holman, S. Seager, J. H. Steffen, J. Tarter, W. F. Welsh, C. Allen, L. A. Buchhave, J. L. Christiansen, B. D. Clarke, J.-M. Désert, M. Endl, D. Fabrycky, F. Fressin, M. Haas, E. Horch, A. Howard, H. Isaacson, H. Kjeldsen, J. Kolodziejczak, C. Kulesa, J. Li, P. Machalek, D. McCarthy, P. MacQueen, S. Meibom, T. Miquel, A. Prsa, S. N. Quinn, E. V. Quintana, D. Ragozzine, W. Sherry, A. Shporer, P. Tenenbaum, G. Torres, J. D. Twicken, J. Van Cleve, and L. Walkowicz. Characteristics of planetary candidates observed by Kepler, II: Analysis of the first four months of data. *ArXiv e-prints*, February 2011. 3

F. Bouchy, S. Udry, M. Mayor, C. Moutou, F. Pont, N. Iribarne, R. da Silva, S. Ilovaisky, D. Queloz, N. C. Santos, D. Ségransan, and S. Zucker. ELODIE metallicity-biased search for transiting Hot Jupiters. II. A very hot Jupiter transiting the bright K star HD 189733. *A&A*, 444:L15–L19, December 2005. doi: 10.1051/0004-6361:200500201. 260

C. J. Burke, P. R. McCullough, L. E. Bergeron, D. Long, R. L. Gilliland, E. P. Nelan, C. M. Johns-Krull, J. A. Valenti, and K. A. Janes. NICMOS Observations of the Transiting Hot Jupiter XO-1b. *ApJ*, 719:1796–1806, August 2010. doi: 10.1088/0004-637X/719/2/1796. 120, 122

R. P. Butler, S. S. Vogt, G. W. Marcy, D. A. Fischer, J. T. Wright, G. W. Henry, G. Laughlin, and J. J. Lissauer. A Neptune-Mass Planet Orbiting the Nearby M Dwarf GJ 436. *ApJ*, 617:580–588, December 2004. doi: 10.1086/425173. 301

C. Cáceres, V. D. Ivanov, D. Minniti, D. Naef, C. Melo, E. Mason, F. Selman, and G. Pietrzynski. High cadence near infrared timing observations of extrasolar planets. I. GJ 436b and XO-1b. *A&A*, 507:481–486, November 2009. doi: 10.1051/0004-6361/200810908. 302, 304

A. Collier Cameron, F. Bouchy, G. Hebrard, P. Maxted, D. Pollacco, F. Pont, I. Skillen, B. Smalley, R. A. Street, R. G. West, D. M. Wilson, S. Aigrain, D. J. Christian, W. I. Clarkson, B. Enoch, A. Evans, A. Fitzsimmons, M. Fleenor, M. Gillon, C. A. Haswell, L. Hebb, C. Hellier, S. T. Hodgkin, K. Horne, J. Irwin, S. R. Kane, F. P. Keenan, B. Loeillet, T. A. Lister, M. Mayor, C. Moutou, A. J. Norton, J. Osborne, N. Parley, D. Queloz, R. Ryans, A. H. M. J. Triaud, S. Udry, and P. J. Wheatley. WASP-1b and WASP-2b: two new transiting

exoplanets detected with superWASP and SOPHIE. *MNRAS*, 375:951–957, 2007. 14, 31

J. A. Carter, J. N. Winn, R. Gilliland, and M. J. Holman. Near-Infrared Transit Photometry of the Exoplanet HD 149026b. *ApJ*, 696:241–253, May 2009. doi: 10.1088/0004-637X/696/1/241. 62

D. Charbonneau, T. M. Brown, R. W. Noyes, and R. L. Gilliland. Detection of an extrasolar planet atmosphere. *ApJ*, 568:377–384, 2002. 45

D. Charbonneau, J. N. Winn, M. E. Everett, D. W. Latham, M. J. Holman, G. A. Esquerdo, and F. T. O'Donovan. Precise Radius Estimates for the Exoplanets WASP-1b and WASP-2b. *ApJ*, 658:1322–1327, April 2007. doi: 10.1086/512008. 232, 241, 286

G. Chauvin, A.-M. Lagrange, C. Dumas, B. Zuckerman, D. Mouillet, I. Song, J.-L. Beuzit, and P. Lowrance. A giant planet candidate near a young brown dwarf. Direct VLT/NACO observations using IR wavefront sensing. *A&A*, 425: L29–L32, 2004. 37

J. L. Christiansen, S. Ballard, D. Charbonneau, N. Madhusudhan, S. Seager, M. J. Holman, D. D. Wellnitz, D. Deming, M. F. A'Hearn, and the EPOXI Team. Studying the atmosphere of the exoplanet HAT-P-7 via secondary eclipse measurements with EPOXI, Spitzer and Kepler. *ApJ*, 710:97–104, 2010. 45, 47

A. Collier Cameron, F. Bouchy, G. Hébrard, P. Maxted, D. Pollacco, F. Pont, I. Skillen, B. Smalley, R. A. Street, R. G. West, D. M. Wilson, S. Aigrain, D. J. Christian, W. I. Clarkson, B. Enoch, A. Evans, A. Fitzsimmons, M. Fleenor, M. Gillon, C. A. Haswell, L. Hebb, C. Hellier, S. T. Hodgkin, K. Horne, J. Irwin, S. R. Kane, F. P. Keenan, B. Loeillet, T. A. Lister, M. Mayor, C. Moutou, A. J. Norton, J. Osborne, N. Parley, D. Queloz, R. Ryans, A. H. M. J. Triaud, S. Udry, and P. J. Wheatley. WASP-1b and WASP-2b: two new transiting exoplanets detected with SuperWASP and SOPHIE. *MNRAS*, 375:951–957, March 2007a. doi: 10.1111/j.1365-2966.2006.11350.x. 216, 217

A. Collier Cameron, D. M. Wilson, R. G. West, L. Hebb, X.-B. Wang, S. Aigrain, F. Bouchy, D. J. Christian, W. I. Clarkson, B. Enoch, M. Esposito, E. Guenther, C. A. Haswell, G. Hébrard, C. Hellier, K. Horne, J. Irwin, S. R. Kane,

B. Loeillet, T. A. Lister, P. Maxted, M. Mayor, C. Moutou, N. Parley, D. Pollacco, F. Pont, D. Queloz, R. Ryans, I. Skillen, R. A. Street, S. Udry, and P. J. Wheatley. Efficient identification of exoplanetary transit candidates from SuperWASP light curves. *MNRAS*, 380:1230–1244, September 2007b. doi: 10.1111/j.1365-2966.2007.12195.x. 217

S. Daemgen, F. Hormuth, W. Brandner, C. Bergfors, M. Janson, S. Hippler, and T. Henning. Binarity of transit host stars. Implications for planetary parameters. *A&A*, 498:567–574, May 2009. doi: 10.1051/0004-6361/200810988. 217

M. Deleuil, H. J. Deeg, R. Alonso, F. Bouchy, D. Rouan, M. Auvergne, A. Baglin, S. Aigrain, J. M. Almenara, M. Barbieri, P. Barge, H. Bruntt, P. Bordé, A. Collier Cameron, Sz. Csizmadia, R. De la Reza, R. Dvorak, A. Erikson, M. Fridlund, D. Gandolfi, M. Gillon, E. Guenther, T. Guillot, A. Hatzes, G. Hébrard, L. Jorda, H. Lammer, A. Léger, A. Llebaria, B. Loeillet, M. Mayor, T. Mazeh, C. Moutou, M. Ollivier, M. Pätzold, F. Pont, D. Queloz, H. Rauer, J. Schneider, A. Shporer, G. Wuchterl, and S. Zucker. Transiting planets from the CoRoT space mission IV. CoRoT-Exo-3b: the first secure inhabitant of the brown-dwarf desert. *A&A*, 491:889–897, 2008. 2

D. Deming, J. Harrington, G. Laughlin, S. Seager, S. B. Navarro, W. C. Bowman, and K. Horning. Spitzer Transit and Secondary Eclipse Photometry of GJ 436b. *ApJ Letters*, 667:L199–L202, October 2007. doi: 10.1086/522496. 302

D. Deming, H. Knutson, E. Agol, J.-M. Desert, A. Burrows, J. J. Fortney, D. Charbonneau, N. B. Cowan, G. Laughlin, J. Langton, A. P. Showman, and N. K. Lewis. Warm-Spitzer photometry of the transiting exoplanets CoRoT-1b and CoRoT-2b at secondary eclipse. *ApJ*, Awaiting Publication, 2010. 47

J.-M. Désert, A. Vidal-Madjar, A. Lecavelier des Etangs, D. Sing, D. Ehrenreich, G. Hébrard, and R. Ferlet. TiO and VO broad band absorption features in the optical spectrum of the atmosphere of the hot-Jupiter HD209458b. *A&A*, 492: 585–592, 2008. 43

J.-M. Désert, A. Lecavelier des Etangs, G. Hébrard, D. K. Sing, D. Ehrenreich, R. Ferlet, and A. Vidal-Madjar. Search for Carbon Monoxide in the Atmosphere

of the Transiting Exoplanet HD 189733b. *ApJ*, 699:478–485, July 2009. doi: 10.1088/0004-637X/699/1/478. 263

V. W. Dixon and et al. *Cosmic Origins Spectrograph Instrument Handbook, Version* 2.0. Baltimore:STScI, 2010. 57

L. Dressel, M. H. Wong, C. Pavlovsky, K. Long, and et al. *Wide Field Camera 3 Instrument Handbook, Version* 2.1. Baltimore:STScI, 2010. 57

D. C. Fabrycky. What to expect from transiting multiplanet systems. In *Transiting Planets. Proceedings IAU Symposium No.*253, 2008. 33

D. A. Fischer, G. W. Marcy, R. P. Butler, S. S. Vogt, G. Laughlin, G. W. Henry, D. Abouav, K. M. G. Peek, J. T. Wright, J. A. Johnson, C. McCarthy, and H. Isaacson. Five planets orbiting 55Cancri. *ApJ*, 675:790–801, 2008. 3

J. J. Fortney, K. Lodders, M. S. Marley, and R. S. Freedman. A unified theory for the atmospheres of the Hot and Very-Hot Jupiters: two classes of irradiated atmospheres. *ApJ*, 678:1419–1435, 2008. 43

L. Fossati, S. Bagnulo, A. Elmasli, C. A. Haswell, S. Holmes, O. Kochukhov, E. L. Shkolnik, D. V. Shulyak, D. Bohlender, B. Albayrak, C. Froning, and L. Hebb. A detailed spectropolarimetric study of the planet-hosting star WASP-12b. *ApJ*, 720:872–886, 2010a. 57

L. Fossati, C. A. Haswell, C. S. Froning, L. Hebb, S. Holmes, U. Kolb, Ch. Helling, A. Carter, P. Wheatley, A. Collier Cameron, B. Loeillet, D. Pollacco, R. Street, H. C. Stempels, E. Simpson, S. Udry, Y. C. Joshi, R. G. West1, I. Skillen, and D. Wilson. Metals in the exosphere of the highly-irradiated planet WASP-12b. *ApJ*, 714:L222–L227, 2010b. 45, 57

N. P. Gibson, F. Pont, and S. Aigrain. A new look at NICMOS transmission spectroscopy of HD 189733, GJ-436 and XO-1: no conclusive evidence for molecular features. *MNRAS*, 411:2199–2213, March 2011. doi: 10.1111/j.1365-2966.2010.17837.x. 62, 262, 299, 300, 304

R. L. Gilliland. Differential NICMOS spectroscopy at high S/N. In *The 2005 HST Calibration Workshop*, 2005. 61, 78, 126

M. Gillon, B.-O. Demory, T. Barman, X. Bonfils, T. Mazeh, F. Pont, S. Udry, M. Mayor, and D. Queloz. Accurate Spitzer infrared radius measurement for the hot Neptune GJ 436b. *A&A*, 471:L51–L54, September 2007a. doi: 10.1051/0004-6361:20078283. 184

M. Gillon, F. Pont, B.-O. Demory, F. Mallmann, M. Mayor, T. Mazeh, D. Queloz, A. Shporer, S. Udry, and C. Vuissoz. Detection of transits of the nearby hot Neptune GJ 436 b. *A&A*, 472:L13–L16, September 2007b. doi: 10.1051/0004-6361:20077799. 302

A. Gould and A. Loeb. Discovering planetary systems through gravitational microlenses. *ApJ*, 396:104–114, September 1992. doi: 10.1086/171700. 36

C. J. Grillmair, D. Charbonneau, A. Burrows, L. Armus, J. Stauffer, V. Meadows, J. Van Cleve, and D. Levine. A Spitzer Spectrum of the Exoplanet HD 189733b. *ApJ Letters*, 658:L115–L118, April 2007. doi: 10.1086/513741. 262

C. J. Grillmair, A. Burrows, D. Charbonneau, L. Armus, J. Stauffer, V. Meadows, J. van Cleve, K. von Braun, and D. Levine. Strong water absorption in the dayside emission spectrum of the planet HD189733b. *Nature*, 456:767–769, December 2008. doi: 10.1038/nature07574. 262

C. A. Haswell. *Transiting Exoplanets.* Cambridge University Press, 2010. 25

A. P. Hatzes, E. W. Guenther, M. Endl, W. D. Cochran, M. P. Döllinger, and A. Bedalov. A giant planet around the massive giant star HD13189. *A&A*, 437:743–751, 2005. 7

L. Hebb, A. Collier-Cameron, A. H. M. J. Triaud, T. A. Lister, B. Smalley, P. F. L. Maxted, C. Hellier, D. R. Anderson, D. Pollacco, M. Gillon, D. Queloz, R. G. West, S. Bentley, B. Enoch, C. A. Haswell, K. Horne, M. Mayor, F. Pepe, D. Segransan, I. Skillen, S. Udry, and P. J. Wheatley. WASP-19b: The shortest period transiting exoplanet yet discovered. *ApJ*, 708:224–231, 2010. 3

G. W. Henry, G. W. Marcy, R. P. Butler, and S. S. Vogt. A transiting "51Peg-like" planet. *ApJ*, 529:L41–L44, 2000. 13

A. Hewish. Pulsars. *Annual Review of Astronomy and Astrophysics*, 8:265, 1970. 40

M. J. Holman, D. C. Fabrycky, D. Ragozzine, E. B. Ford, J. H. Steffen, W. F. Welsh, J. J. Lissauer, D. W. Latham, G. W. Marcy, L. M. Walkowicz, N. M. Batalhaand J. M. Jenkins, J. F. Rowe, W. D. Cochran, F. Fressin, G. Torres, L. A. Buchhave, D. D. Sasselov, W. J. Borucki, D. G. Koch, G. Basri, T. M. Brown, D. A. Caldwell, D. Charbonneau, E. W. Dunham, T. N. Gautier III, J. C. Geary, R. L. Gilliland, M. R. Haas, S. B. Howell, D. R. Ciardi, M. Endl, D. Fischer, G. Frész, J. D. Hartman, H. Isaacson, J. A. Johnson, P. J. MacQueen, A. V. Moorhead, R. C. Morehead, and J. A. Orosz. Kepler-9: A system of multiple planets transiting a sun-like star confirmed by timing variations. *Science*, 330:51–54, 2010. 3

HubbleSite. `http://hubblesite.org/`, 2010. 53

P. Kalas, J. R. Graham, E. Chiang, M. P. Fitzgerald, M. Clampin, E. S. Kite, K. Stapelfeldt, C. Marois, and J. Krist. Optical images of an exosolar planet 25 light-years from Earth. *Science*, 322:1345–1348, 2008. 37, 38, 50

H. A. Knutson, D. Charbonneau, L. E. Allen, J. J. Fortney, E. Agol, N. B. Cowan, A. P. Showman, C. S. Cooper, and S. T. Megeath. A map of the day-night contrast of the extrasolar planet HD189733b. *Nature*, 447:183–186, 2007. 48

M. Konacki, G. Torres, D. D. Sasselov, and S. Jha. A transiting extrasolar giant planet around the star OGLE-TR-10. *ApJ*, 624:372–377, 2005. 5

J. J. Lissauer, D. C. Fabrycky, E. B. Ford, W. J. Borucki, F. Fressin, G. W. Marcy, J. A. Orosz, J. F. Rowe, G. Torres, W. F. Welsh, N. M. Batalha, S. T. Bryson, L. A. Buchhave, D. A. Caldwell, J. A. Carter, D. Charbonneau, J. L. Christiansen, W. D. Cochran, J.-M. Desert, E. W. Dunham, M. N. Fanelli, J. J. Fortney, T. N. Gautier, III, J. C. Geary, R. L. Gilliland, M. R. Haas, J. R. Hall, M. J. Holman, D. G. Koch, D. W. Latham, E. Lopez, S. McCauliff, N. Miller, R. C. Morehead, E. V. Quintana, D. Ragozzine, D. Sasselov, D. R. Short, and J. H. Steffen. A closely packed system of low-mass, low-density planets transiting Kepler-11. *Nature*, 470:53–58, February 2011. doi: 10.1038/nature09760. 13

G. Maciejewski, R. Neuhauser, R. Errmann, M. Mugrauer, Ch. Adam, A. Berndt, T. Eisenbeiss, S. Fiedler, Ch. Ginski, M. Hohle, U. Kramm, C. Marka,

M. Moualla, T. Pribulla, St. Raetz, T. Roell, T. O. B. Schmidt, M. Seeliger, I. Spaleniak, N. Tetzlaff, and L. Trepl. Towards the Rosetta Stone of planet formation. In *Detection and Dynamics of Transiting Exoplanets*, 2010a. 2

G. Maciejewski, D. Dimitrov, R. Neuhauser, A. Niedzielski, St. Raetz, Ch. Ginski, Ch. Adam, C. Marka, M. Moualla, and M. Mugrauer. Transit timing variation in exoplanet WASP-3b. *MNRAS*, 407:2625–2631, 2010b. 33, 34

N. Madhusudhan and S. Seager. High Metallicity and Non-equilibrium Chemistry in the Dayside Atmosphere of hot-Neptune GJ 436b. *ApJ*, 729:41–+, March 2011. doi: 10.1088/0004-637X/729/1/41. 304

K. Mandel and E. Agol. Analytic Light Curves for Planetary Transit Searches. *ApJ Letters*, 580:L171–L175, December 2002. doi: 10.1086/345520. 25, 144

A. M. Mandell, L. Drake Deming, G. A. Blake, H. A. Knutson, M. J. Mumma, G. L. Villanueva, and C. Salyk. Non-detection of L-band Line Emission from the Exoplanet HD189733b. *ApJ*, 728:18–+, February 2011. doi: 10.1088/0004-637X/728/1/18. 263, 299

G. Mandushev, F. T. O'Donovan, D. Charbonneau, G. Torres, D. W. Latham, G. Á. Bakos, E. W. Dunham, A. Sozzetti, J. M. Fernández, G. A. Esquerdo, M. E. Everett, T. M. Brown, M. Rabus, J. A. Belmonte, and L. A. Hillenbrand. TrES-4: A transiting Hot Jupiter with very low density. *ApJ*, 667:L195–L198, 2007. 2

C. Marois, B. Macintosh, T. Barman, B. Zuckerman, I. Song, J. Patience, D. Lafreniére, and R. Doyon. Direct imaging of multiple planets orbiting the star HR8799. *Science*, 322:1348–1352, 2008. 3, 37

C. Marois, B. Zuckerman, Q. M. Konopacky, B. Macintosh, and T. Barman. Images of a fourth planet orbiting HR 8799. *Nature*, 468:1080–1083, December 2010. doi: 10.1038/nature09684. 37

A. Maybhate and et al. *ACS Instrument Handbook, Version* 9.0. Baltimore:STScI, 2009. 54

M. Mayor and D. Queloz. A Jupiter-mass planet around a solar-type star. *Nature*, 378:355–359, 1995. 2, 4

M. Mayor, S. Udry, C. Lovis, F. Pepe, D. Queloz, W. Benz, J.-L. Bertaux, F. Bouchy, C. Mordasini, , and D. Segransan. The HARPS search for southern exoplanets XIII. a planetary system with 3 super-Earths (4.2, 6.9 and $9.2m_{\oplus}$). *A&A*, 493:639–644, 2008. 3

M. Mayor, X. Bonfils, T. Forveille, X. Delfosse, S. Udry, J.-L. Bertaux, H. Beust, F. Bouchy, C. Lovis, F. Pepe, C. Perrier, D. Queloz, and N. C. Santos. The HARPS search for southern extrasolar planets XVIII. an Earth mass planet in the GJ581 planetary system. *A&A*, 507:487–494, 2009. 3

M. McMaster, J. Biretta, and et al. *WFPC2 Instrument Handbook, Version* 10.0. Baltimore:STScI, 2008. 55

C. Mordasini, Y. Alibert, W. Benz, and D. Naef. Extrasolar planet population synthesis II. statistical comparison with observations. *A&A*, 501:1161–1184, 2009. 2

A. Nota and et al. *FOC Instrument Handbook, Version* 7.0. Baltimore:STScI, 1996. 55

F. Pont, H. Knutson, R. L. Gilliland, C. Moutou, and D. Charbonneau. Detection of atmospheric haze on an extrasolar planet: the 0.55-1.05 μm transmission spectrum of HD 189733b with the HubbleSpaceTelescope. *MNRAS*, 385:109–118, March 2008. doi: 10.1111/j.1365-2966.2008.12852.x. 262, 263, 293

F. Pont, R. L. Gilliland, H. Knutson, M. Holman, and D. Charbonneau. Transit infrared spectroscopy of the hot Neptune around GJ 436 with the Hubble Space Telescope. *MNRAS*, 393:L6–L10, February 2009. doi: 10.1111/j.1745-3933.2008.00582.x. 62, 304

W. H. Press, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipies*, chapter 14. Cambridge University Press, third edition, 2007a. 145, 157

W. H. Press, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipies*, chapter 15. Cambridge University Press, third edition, 2007b. 150, 151, 155, 166

W. H. Press, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipies*, chapter 16. Cambridge University Press, third edition, 2007c. 163, 173

C. Proffitt and et al. *STIS Instrument Handbook, Version* 9.0. Baltimore:STScI, 2010. 54

S. Redfield, M. Endl, W. D. Cochran, and L. Koesterke. Sodium Absorption from the Exoplanetary Atmosphere of HD 189733b Detected in the Optical Transmission Spectrum. *ApJ Letters*, 673:L87–L90, January 2008. doi: 10.1086/527475. 45, 262

I. Ribas and J. Miralda-Escudé. The eccentricity-mass distribution of exoplanets: signatures of different formation mechanisms? *A&A*, 464:779–785, 2007. 49

I. Ribas, A. Font-Ribera, and J.-P. Beaulieu. A $\sim 5M_E$ Super-Earth Orbiting GJ 436? The Power of Near-Grazing Transits. *ApJ Letters*, 677:L59–L62, April 2008. doi: 10.1086/587961. 304

D. D. Sasselov. Extrasolar planets. *nature*, 451:29–31, 2008. 4, 42

J. Schneider. The Extrasolar Planets Encyclopaedia. `http://exoplanet.eu/`, 2010. 2, 3, 4, 31, 36, 37

D. K. Sing, J.-M. Désert, A. Lecavelier Des Etangs, G. E. Ballester, A. Vidal-Madjar, V. Parmentier, G. Hebrard, and G. W. Henry. Transit spectrophotometry of the exoplanet HD 189733b. I. Searching for water but finding haze with HST NICMOS. *A&A*, 505:891–899, October 2009. doi: 10.1051/0004-6361/200912776. 263, 287, 291, 294, 296, 299

D. K. Sing, F. Pont, S. Aigrain, D. Charbonneau, J. . Desert, N. Gibson, R. Gilliland, W. Hayek, G. Henry, H. Knutson, A. Lecavelier des Etangs, T. Mazeh, and L. Tal-Or. Hubble Space Telescope Transmission Spectroscopy of the Exoplanet HD 189733b: High-altitude atmospheric haze in the optical and near-UV with STIS. *ArXiv e-prints*, February 2011. 263

C. J. Skinner, L. E. Bergeron, A. B. Schultz, J. W. MacKenty, A. Storrs, W. Freudling, D. Axon, H. Bushouse, D. Calzetti, L. Colina, D. Daou, D. Gilmore, S. T. Holfeltz, J. Najita, K. Noll, C. E. Ritchie, W. B. Sparks, and A. Suchkov. On-orbit properties of the NICMOS detectors on HST. In A. M. Fowler, editor, *Society of Photo-Optical Instrumentation Engineers*

(SPIE) Conference Series, volume 3354 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 2–13, August 1998. 298

I. A. G. Snellen, S. Albrecht, E. J. W. de Mooij, and R. S. Le Poole. Ground-based detection of sodium in the transmission spectrum of exoplanet HD209458b. *A&A*, 487:357–362, 2008. 45

D. R. Soderblom, A. Gonnella, S. J. Hulbert, C. Leitherer, A. Schultz, and L. E. Sherbert. *GHRS Instrument Handbook, Version* 6.0. Baltimore:STScI, 1995. 56

J. Southworth. Homogeneous studies of transiting extrasolar planets - III. Additional planets and stellar models. *MNRAS*, 408:1689–1713, November 2010. doi: 10.1111/j.1365-2966.2010.17231.x. 217, 260, 302, 315

D. S. Spiegel, A. Burrows, and J. A. Milsom. The deuterium-burning mass limit for brown dwarfs and giant planets. *ApJ*, Awaiting Publication, 2010. 30

K. B. Stevenson, J. Harrington, S. Nymeyer, N. Madhusudhan, S. Seager, W. C. Bowman, R. A. Hardy, D. Deming, E. Rauscher, and N. B. Lust. Possible thermochemical disequilibrium in the atmosphere of the exoplanet GJ 436b. *Nature*, 464:1161–1164, April 2010. doi: 10.1038/nature09013. 304

STScI. Space telescope science institute, johns hopkins university. `http://www.stsci.edu/`, 2010. 52, 73

M. R. Swain, G. Vasisht, and G. Tinetti. The presence of methane in the atmosphere of an extrasolar planet. *Nature*, 452:329–331, 2008. 45, 61, 259, 262, 263, 274, 287, 291, 294, 296, 297, 298, 299, 300

M. R. Swain, G. Tinetti, G. Vasisht, P. Deroo, C. Griffith, J. Bouwman, P. Chen, Y. Yung, A. Burrows, L. R. Brown, J. Matthews, J. F. Rowe, R. Kuschnig, and D. Angerhausen. Water, Methane, and Carbon Dioxide Present in the Dayside Spectrum of the Exoplanet HD 209458b. *ApJ*, 704:1616–1621, October 2009a. doi: 10.1088/0004-637X/704/2/1616. 62

M. R. Swain, G. Vasisht, G. Tinetti, J. Bouwman, P. Chen, Y. Yung, D. Deming, and P. Deroo. Molecular Signatures in the Near-Infrared Dayside Spectrum of HD 189733b. *ApJ Latters*, 690:L114–L117, January 2009b. doi: 10.1088/0004-637X/690/2/L114. 262, 273

M. R. Swain, G. Vasisht, G. Tinetti, J. Bouwman, P. Chen, Y. Yung, D. Deming, and P. Deroo. Molecular Signatures in the Near-Infrared Dayside Spectrum of HD 189733b. *ApJ Letters*, 690:L114–L117, January 2009c. doi: 10.1088/0004-637X/690/2/L114. 61

M. R. Swain, P. Deroo, C. A. Griffith, G. Tinetti, A. Thatte, G. Vasisht, P. Chen, J. Bouwman, I. J. Crossfield, D. Angerhausen, C. Afonso, and T. Henning. A ground-based near-infrared emission spectrum of the exoplanet HD189733b. *Nature*, 463:637–639, February 2010. doi: 10.1038/nature08775. 262, 263

D. Thatte, T. Dahlen, and et al. *NICMOS Data Handbook, Version* 8.0. Baltimore:STScI, 2009. 77

G. Tinetti, A. Vidal-Madjar, M.-C. Liang, J.-P. Beaulieu, Y. Yung, S. Carey, R. J. Barber, J. Tennyson, I. Ribas, N. Allard, G. E. Ballester, D. K. Sing, and F. Selsis. Water vapour in the atmosphere of a transiting extrasolar planet. *Nature*, 448:169–171, July 2007. doi: 10.1038/nature06002. 262

G. Tinetti, P. Deroo, M. R. Swain, C. A. Griffith, G. Vasisht, L. R. Brown, C. Burke, and P. McCullough. Probing the Terminator Region Atmosphere of the Hot-Jupiter XO-1b with Transmission Spectroscopy. *ApJ Letters*, 712: L139–L142, April 2010. doi: 10.1088/2041-8205/712/2/L139. 45, 62

K. Todorov, D. Deming, J. Harrington, K. B. Stevenson, W. C. Bowman, S. Nymeyer, J. J. Fortney, and G. A. Bakos. Spitzer IRAC secondary eclipse photometry of the transiting extrasolar planet HAT-P-1b. *ApJ*, 708:498–504, 2010. 47

G. Torres. The Transiting Exoplanet Host Star GJ 436: A Test of Stellar Evolution Models in the Lower Main Sequence, and Revised Planetary Parameters. *ApJ Letters*, 671:L65–L68, December 2007. doi: 10.1086/524886. 302

A. H. M. J. Triaud, A. Collier Cameron, D. Queloz, D. R. Anderson, M. Gillon, L. Hebb, C. Hellier, B. Loeillet, P. F. L. Maxted, M. Mayor, F. Pepe, D. Pollacco, D. Ségransan, B. Smalley, S. Udry, R. G. West, and P. J. Wheatley. Spin-orbit angle measurements for six southern transiting planets. New insights into the dynamical origins of hot Jupiters. *A&A*, 524:A25+, December 2010. doi: 10.1051/0004-6361/201014525. 217, 260

A. Udalski, B. Paczynski, K. Zebrun, M. Szymanski, M. Kubiak, I. Soszynski, O. Szewczyk, L. Wyrzykowski, and G. Pietrzynski. The optical gravitational lensing experiment. search for planetary and low-luminosity object transits in the galactic disk. Results of 2001 campaign. *Acta Astronomica*, 52:1–37, 2002. 13

A. Viana, T. Wiklind, and et al. *NICMOS Instrument Handbook, Version* 11.0. Baltimore:STScI, 2009. 56, 58, 117, 205

A. Vidal-Madjar, A. Lecavelier des Etangs, J.-M. Désert, G. E. Ballester, R. Ferlet, G. Hébrard, and M. Mayor. An extended upper atmosphere around the extrasolar planet HD209458b. *Nature*, 422:143–146, March 2003. doi: 10.1038/nature01448. 45, 54

A. Wolszczan and D. A. Frail. A planetary system around the millisecond pulsar PSR1257 + 12. *Nature*, 355:145–147, 1992. 1, 7, 40

A. N. Youdin and F. H. Shu. Planetesimal formation by gravitational instability. *ApJ*, 580:494–505, 2002. 50